

Tue Sep 21 09:16:38 2004

us-10-009-945-4.rspt

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:36:41; Search time 91.0211 Seconds  
(without alignments)  
2592.890 Million cell updates/sec

Title: US-10-009-945-4  
Perfect score: 4038  
Sequence: 1 MSNPGRRNGPYKRLTYLC.....EKLYEKLLTAIEETGFAVE 748

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rv1rus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3041	75.3	728	11	Q8K300
2	2313	57.3	1061	5	Q3V853
3	1910	47.3	376	13	Q98RS4
4	1668.5	41.3	355	11	Q8BSC0
5	1552	38.4	786	3	Q14326
6	1418	35.1	911	4	Q8N5A7
7	1398	34.6	955	4	Q96PU5
8	1386	34.3	947	4	Q7Z5F1
9	1384	34.2	295	4	Q8NDP8
10	1379	34.1	258	11	Q9C8S3
11	1378	34.0	975	4	Q7Z5N3
12	1371	33.9	967	4	Q7Z5F2
13	1369	33.8	820	4	Q9NT88
14	1364.5	33.6	855	11	Q8CFH0
15	1357.5	33.5	854	4	Q9H2W4
16	1354.5	33.5	854	4	Q9H2W4

17	1354	33.5	835	11	Q8BRT9	Q8BRT9 mus musculu
18	1353.5	33.5	995	4	Q43165	Q43165 homo sapien
19	1353	33.5	855	4	Q8WUJ9	Q8WUJ9 homo sapien
20	1353	33.5	858	4	Q9BM58	Q9BM58 homo sapien
21	1346.5	33.3	855	11	Q99PK2	Q99PK2 mus musculu
22	1345.5	33.3	971	13	Q42573	Q42573 xenopus lae
23	1341.5	33.2	834	5	Q95RE4	Q95RE4 drosophila
24	1333	33.0	945	5	Q9YOH4	Q9YOH4 drosophila
25	1309.5	32.3	956	5	Q8IOR6	Q8IOR6 drosophila
26	1297	32.1	1007	5	Q9VUJ3	Q9VUJ3 drosophila
27	1274.5	31.6	792	5	Q95XU3	Q95XU3 caenorhabdi
28	1271.5	31.5	794	5	Q9N2T7	Q9N2T7 caenorhabdi
29	1244	30.8	728	4	Q9BKW4	Q9BKW4 caenorhabdi
30	1224.5	30.3	834	5	Q9STQ0	Q9STQ0 drosophila
31	1220.5	30.2	518	5	Q8T0C8	Q8T0C8 drosophila
32	1092	27.0	1581	4	Q9P2P5	Q9P2P5 homo sapien
33	1050	26.0	1585	4	Q9HCT7	Q9HCT7 homo sapien
34	1046.5	25.9	671	3	Q9UTG2	Q9UTG2 schizosacch
35	1032	25.6	222	11	Q80TB1	Q80TB1 mus musculu
36	1011.5	25.0	1583	11	Q8K4P8	Q8K4P8 mus musculu
37	926.5	22.9	437	5	Q9W325	Q9W325 drosophila
38	877.5	21.7	339	11	Q8BIA6	Q8BIA6 mus musculu
39	859	21.3	403	11	Q8VDI4	Q8VDI4 mus musculu
40	859	21.3	444	11	Q7TWI8	Q7TWI8 mus musculu
41	859	21.3	1080	11	Q92IM5	Q92IM5 mus musculu
42	859	21.3	1484	11	Q8CFH2	Q8CFH2 mus musculu
43	859	21.3	1906	4	Q15029	Q15029 homo sapien
44	859	21.3	3360	4	Q8NG67	Q8NG67 homo sapien
45	858	21.2	3227	3	Q13834	Q13834 schizosacch

## ALIGNMENTS

RESULT 1  
ID Q8K300 PRELIMINARY; PRT; 728 AA.  
AC Q8K300;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to E3 ubiquitin ligase SMURF1.  
GN 493043310R1K.  
GN Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
DR EMBL; BC029097; AAH29097.1; -;  
DR MGD; MGI:1923038; 493043310R1K.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:001674; F:ligase activity; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CaLb.  
DR InterPro; IPR000569; HECH domain.  
DR InterPro; IPR01202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECH; 1.  
DR Pfam; PF00357; WW; 2.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECH; 1.  
DR SMART; SM00456; WW; 2.  
DR PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS02037; HECH; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_2; 1.  
DR PROSITE; PS50020; WW\_DOMAIN\_1; 2.  
DR PROSITE; PS50020; WW\_DOMAIN\_2; 2.

KW L1gase. 728 AA; 83083 MW; C355291B9D8AD757 CRC64;  
 SEQ SEQUENCE

Query Match 75.3%; Score 3041; DB 11; Length 728;  
 Best Local Similarity 75.0%; Pred. No. 1.7e-229;  
 Matches 572; Conservative 63; Mismatches 78; Indels 50; Gaps 7;

QY 1 MNNPERRRNG-FVKRLRLTLCAKAKLVKDFPLPDPFAKVVVDGSGGCHSDTDYKNTLP 59  
 DB 1 MNNPGRNRSSIKRLRLTLCAKAKLVKDFPLPDPFAKVVVDGSGGCHSDTDYKNTLP 60  
 QY 60 KNNQHYDLYIGKSDSVTISVMNKKIKKKQAGAGFGICRLSNAINRLKDDGYORLPDCK 119  
 DB 61 KNSQHYDLYIGKSDSVTISVMNKKIKKKQAGAGFGICRLSNAINRLKDDGYORLPDCK 120  
 QY 120 LCPNNDYRGQIVVLSGSRDRIIGTGGQVDCSRIPNDLPDGMEEERTASGRIOYLNHI 179  
 DB 121 INPSDLDVARGQIVVLSGSRDRIIGTGGQVDCSRIPNDLPDGMEEERTASGRIOYLNHI 163  
 QY 180 TTTTQWERPTRASYSFSGFPLSCFVDPENTPISGTNGATG-----QSSDPLAE 230  
 DB 164 -----EDSGRGRPLSLMEEPAPYIDGTAAAGGNCRFVESPQDQRLV 209  
 QY 221 RRVRSQRHNVN-----SRTHLTPPDLEGEYORITTOGGQVYPLHTQGVSTWHDPRVPR 286  
 DB 210 QRLRNPVYRGFLQTPQNRPHGHQSPBELBEGYEORTTVQGVYFLHTQGVSTWHDPRIPR 269  
 QY 287 DLSNICEIGPLPGWEIRNTANCGRVYFVDHNNRTQFTTPR.SANLHLVLNRQNLKD 346  
 DB 270 DLSNICEIGPLPGWEIRNTANCGRVYFVDHNNRTQFTTPR---LHHIMHQCOLKE 325  
 QY 347 QQQQ-QVAVSLCPDDTECTVPRYKRDLYQKLIROELSDQOQOAGCHRIEVSREIFEE 405  
 DB 326 PSQPLQPSSEGSVEDELPQORYERDLYQKLVRLHLSLQDQAGHCRLEVSREIFEE 385  
 QY 406 SYRQVMMKRPDMLKRLMKIRGREGGLDYGVAEMLYLSHEMLNRYGIFQYGRDDYI 465  
 DB 386 SYRQVMMKRPDMLKRLMKIRGREGGLDYGVAEMLYLSHEMLNRYGIFQYGRDDYI 445  
 QY 466 TLQINPDSAVNPEHISYFPEFVGRIMGAVFPGHYIDGFTLPFKYKOLGKISITLDDMELV 525  
 DB 446 TLQINPDSINPDHLSYFPGVGRIMGAVFPGHYIDGFTLPFKYKOLGKISITLDDMELV 505  
 QY 526 DPLHNSLWILNDITGVLDHTFCVHNAYGEIIOHELKPKNGKSIIPVNEENKKEYVRLY 585  
 DB 506 DPLHNSLWILNDITGVLDHTFCVHNAYGEIIOHELKPKNGKSIIPVNEENKKEYVRLY 565  
 QY 586 VNMRFELRGIEAQFALQKGFENEVTPQHLKTFDEKELIILICGLKIDVNDKVNTRLKH 645  
 DB 566 VNMRFELRGIEAQFALQKGFENEVTPQHLKTFDEKELIILICGLKIDVNDKVNTRLKH 625  
 QY 646 CTSPDSNIYKFWKVAVEFFDEERRARLLQFTVGSRRVPLQGFALQGAAGRLFTIHQIDA 705  
 DB 626 CVADSNIVRMFWQAVETFEDEERRARLLQFTVGSRRVPLQGFALQGAAGRLFTIHQIDA 685  
 QY 706 CTNNLPKRAHGTGNRIDIPIYSEYKYEKILLTAEEFCGAV 748  
 DB 686 NTDNLPKRAHGTGNRIDIPIYSEYKYEKILLTAEEFCGAV 728

RESULT 2  
 Q9V853 PRELIMINARY; PRT; 1061 AA.  
 AC Q9V853: Q9U8W2;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE CG4943 protein (E3 ubiquitin ligase) (Ubiquitin-protein ligase) (Smad-ubiquitin E3 ligase Smurf1).  
 GN IACK OR SMURF OR SMURF1 OS CG4943.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA Gordon R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,  
 RA Anelli J.F., Agbayani A., An H.-U., Andrews-Frankoon C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
 RA Borokva D., Botchan M.R., Bouck R., Brockstein P., Broctier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,  
 RA Iasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Izu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merklov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.G.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Snue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E.C., Spieding A.C., Stapleton M., Strong R., Sun E.,  
 RA Switzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Massaman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banton J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreira S., Frise E., Galie R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Paclet B., Paragass V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Switzkas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
 RT "Sequencing of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP Mistra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Seattle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
 RA "Annotation of Drosophila melanogaster genome.";  
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.  
 RA Adams M.D., Colniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RU Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Flybase;  
 RU Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Laurenson A., Hawley S.;  
 RU Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Podos S.D., Hanson K.K., Wang Y.-C., Ferguson E.L.;  
 RU "The Dmurt ubiquitin-protein ligase restricts BMP signaling spatially and temporally during Drosophila development.";  
 RL Dev. Cell 1:0-0 (2001).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RA Liang Y.-Y., Lin X., Feng X.-H.;  
 RU "dsurf1, a Smad-ubiquitin E3 ligase, specifically targets dpp-activated Mad protein for degradation.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
 DR EMBL: AEO03802; AAF57824.3; -;  
 DR EMBL: AF216521; AAF21125.1; -;  
 DR EMBL: AF416571; AAL09691.1; -;  
 DR EMBL: AF464851; AAM09646.1; -;  
 DR HSSP: Q13526; 1PIN.  
 DR Flybase: FBgn0029006; Jack.  
 DR GO: GO:0005622; C:intracellular; IEA.  
 DR GO: GO:0016874; E:ligase activity; IEA.  
 DR GO: GO:0004942; F:ubiquitin-protein ligase activity; IEA.  
 DR GO: GO:0006512; F:ubiquitin cycle; IEA.  
 DR InterPro: IPR000973; C2 CALB.  
 DR InterPro: IPR000569; HECT domain.  
 DR InterPro: IPR001202; ww\_Rsp5\_wmp.  
 DR Pfam: PF00168; C2; 1.  
 DR Pfam: PF00632; HECT; 1.  
 DR Pfam: PF00397; WW; 3.  
 DR SMART: SMO0456; WW; 3.  
 DR SMART: SMO0456; WW; 3.  
 DR PROSITE: PSS00499; C2 DOMAIN 1; 1.  
 DR PROSITE: PSS00064; C2 DOMAIN 2; 1.  
 DR PROSITE: PSS0237; HECT; 1.  
 DR PROSITE: PSS0159; WW DOMAIN 1; 2.  
 DR PROSITE: PSS0020; WW DOMAIN 2; 3.  
 DR LIGASE.  
 KM  
 SQ SEQUENCE 1061 AA; 115675 MW; 68BCC50F5129163 CRC64;

Query Match 57.3%; Score 2313; DB 5; Length 1061;  
 Best Local Similarity 45.1%; Pred. No. 3.5e-172;  
 Matches 476; Conservative 107; Mismatches 157; Indels 316; Gaps 15;

DB 246 HRSRDLVTASDERRRSTELLSSVGKENTSPPTPVGATTPGKKTSSNSAGRTLEQ 305  
 QY 207 -----DENTPIGCT-----N 216  
 DB 306 RPTNEPATPTSTTSASVRLHSDNDHVKTPKQINGHAPPESTPTSGQYVNGNAQN 365  
 QY 217 GAT-----CGQSSDPRLAER----- 231  
 DB 366 GSTISGSGGQAQPOSASNWTQEDAAITTSPTSTTSPPRHSGSPPTPNISPPASVPSA 425  
 QY 232 -----RVRSQRHNYMSRTHLTTP----- 251  
 DB 426 NGNVHSPNANSTPAGSGGSRSYTATPGQSRQSRSSRQGESESTARRSSRGTNGTS 485  
 QY 252 -----DIPGEQRTTGGQVYPLHTQTVSTWHDPR 283  
 DB 486 GGGGGGGGQRYASAAIAANQAPPLDIPGTEKMTTQGGQVYFHTIGVSTWHDPR 545  
 QY 284 VPRDL--SNINCEELGPLPGMEIRNTATGRVYFVDHNNRTTQTPDRLSANLHLVNRQ 341  
 DB 546 IPRPDYQHLTLDAIGLPGMEQRKTAAGRVYVDHNNRTTQTPDRLSGSLIQMIRRG 605  
 QY 342 N-----QKDCQQQQQVSL-----CPDTE 361  
 DB 606 TVPPTSAANAGTPAPSPATPATPSAAAAVPPQATPASNAATPTTLTTNPPHRIVPLPQ 665  
 QY 362 CL-----TVPRYKADIVQKTLIRGEISQOOPQAGHCRIEVSREIFEESYRQYMKRKPX 416  
 DB 666 GLRGADILPKYKRDVNGKRALMTETLQTMQPSGCHRLVSRREIFEESYRLMKRAX 725  
 QY 417 DLMKRLMKFGESEGLDYGVAASMYLISHMLNPPYGLFQYSRDDITYLQINPDNAV 476  
 DB 726 DMKRLMKVFKGESEGLDYGVAASMYLISHMLNPPYGLFQYSRDDITYLQINPDNAV 785  
 QY 477 PEHLSTYFHFVGRINGMAVFGHYIDGGFTLPFYQLLGKSTLIDMELVDPDLNLSYVI 536  
 DB 786 PDHLSYFHFVGRINGMAVFGHYIDGGFTLPFYQLLGKSTLIDMELVDPDLNLSYVI 845  
 QY 537 LENDITGVLDHTPFCEHNAAGEIIQHELKPNKGSIPVNEENKKEVYVLYNMRFLGIEA 596  
 DB 846 LENSISGIIESTFVSENNISFGALVYVHLEKPGASIPVTEENKREYKLYNVMRFGIEQ 905  
 QY 597 QFLALQKGNBYIPOHLLKTDKEXLELIICGLKIDVDMKNAKRLKHCIPDSNIVKWF 656  
 DB 906 QFLALQKGFCELIPIHLLRPDERLELVIGISISIDVDMKNNRRLKHCINETTVQVLMF 965  
 QY 657 WKAVEFPDEERRARLLQFVTGSSRYPLQGFALQ---GAAGPRLFTTH-QIDACTNNLPK 712  
 DB 966 WQVYESISEMRAKLLQFVTGSSRYPLQGFALQSTAVAPRLFTTHLTADVPLQNLPK 1025  
 QY 713 AATCFNRIDIPPESEYKLYELKLTALIEETGFAVE 748  
 DB 1026 AATCFNRIDIPPEYTYQLCDLTLQAVEETGFAVE 1061

RESULT 3  
 Q98TS4 PRELIMINARY; PRT; 376 AA.  
 AC Q98TS4;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE E3 ubiquitin ligase smurf2 (Fragment).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 NC NCBL\_Taxid=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21107656; PubMed=11558580;  
 RA Zhang Y., Chang C., Gehling D.O., Hemmati-Brivanlou A., Derynck R.;

RT "Regulation of Smad degradation and activity by Smurf2, an E3  
 RT ubiquitin ligase.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:974-979(2001).  
 CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
 DR EMBL: AY014181; AAG50422.1; -.  
 DR HSSP: Q13526; 1PIN.  
 DR GO: GO:0016874; F:ligase activity; IEA.  
 DR InterPro: IPR000008; C2.  
 DR InterPro: IPR008973; C2\_GaIB.  
 DR InterPro: IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam: PF00168; C2\_1.  
 DR Pfam: PF00397; WW; 3.  
 DR SMART: SM00239; C2; 1.  
 DR SMART: SM00456; WW; 3.  
 DR PROSITE: PS00004; C2 DOMAIN 2; 1.  
 DR PROSITE: PS01159; WW DOMAIN 1; 1.  
 DR PROSITE: PS0020; WW\_DOMAIN 2; 3.  
 DR LIGase.  
 KW NON\_TER  
 FT SEQUENCE 376 AA; 42652 MW; CA6AB2B5F4C20F98 CRC64;  
 Query Match 47.3%; Score 1910; DB 13; Length 376;  
 Best Local Similarity 94.1%; Pred. No. 3e-141;  
 Matches 354; Conservative 10; Mismatches 8; Indels 4; Gaps 3;  
 QY 1 MSNPRRRNGPVKRLTVLCAKNIWKDFR-LPPFAKVVVDGSGGCHSTDFYKNTLP 59  
 DB 1 MSNQSRRNGPVKRLTVLCAKNIWKDFRGLTGLDSPAFAVVVDGSGGCHSTDFYKNTLP 60  
 QY 60 KMNQHYDLYIKSSSVTSVANKKIKHKQAGLGCVRLLSNAINFLKDTGYQRDLCK 119  
 DB 61 KMNQHYDLYIKSSSVTSVANKKIKHKQAGLGCVRLLSNAINFLKDTGYQRDLCK 120  
 QY 120 LGPNDNTPVRCQIVYSLOSRIQIGTGGVVDGSLFPNDLPDGEERRTASGRIQVYLNH 179  
 DB 121 LGPNDNTPVRCQIVYSLOSRIQIGTGGVVDGSLFPNDLPDGEERRTASGRIQVYLNH 180  
 QY 180 TRTQWBERPVPASVYSSPGRLSCFVDENTPISTINGATCGOSSDRLAERVRSGRR 239  
 DB 181 TRTQWBERPVPASVYSSPGRLSCFVDENTPISTINGATCGOSSDRLAERVRSGRR 240  
 QY 240 NYMSRTHLHPDPDPEGEYEQRTQGGVYFLHTQTVSTWHDPRVPDLNINCEBGL 299  
 DB 241 NYMSRTHLHPDPDPEGEYEQRTQGGVYFLHTQTVSTWHDPRVPDLNINCEBGL 300  
 QY 300 PPGAEINNTATGRVYFVDHNNRTQFTDPRLSANLHLVLRN-OLKDCQGGQVVSIC-- 356  
 DB 301 PPGAEINNTATGRVYFVDHNNRTQFTDPRLSANLHLVLRN-OLKDCQGGQVVSIC-- 360  
 QY 357 PDDEECUTVPRYKDL 372  
 DB 361 PDEVECLTVPRYKDL 376  
 RESULT 4  
 Q8BSC0 ID Q8BSC0 PRELIMINARY; PRT; 355 AA.  
 AC Q8BSC0;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Hypothetical HECT domain (Fragment).  
 GN 4930431E10RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CS7BL/6J; TISSUE=Embryo;  
 RX MEDLINE=2354683; PubMed=12466851;  
 RA THE PANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL: AK034736; BAC28813.1; -.  
 DR MGD: MGI:1923038; 4930431E10RIK.  
 DR GO: GO:0005622; C:intracellular; IEA.  
 DR GO: GO:0004642; F:ubiquitin-protein ligase activity; IEA.  
 DR GO: GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro: IPR000569; HECT domain.  
 DR Pfam: PF00632; HECT; 1.  
 DR SMART: SM00119; HECTC; 1.  
 DR PROSITE: PS00237; HECT; 1.  
 DR Hypothetical protein.  
 KW NON\_TER  
 FT SEQUENCE 355 AA; 41315 MW; 93B39B4C82F86DD1 CRC64;  
 Query Match 41.3%; Score 1668.5; DB 11; Length 355;  
 Best Local Similarity 86.5%; Pred. No. 2.4e-122;  
 Matches 307; Conservative 25; Mismatches 20; Indels 3; Gaps 1;  
 QY 397 VRESEIFESRYQVWKMRPKDIWKRLMKFRGEGLDYGVARWLYLISHMLNPPYGL 456  
 DB 1 VRESEIFESRYQVWKMRPKDIWKRLMKFRGEGLDYGVARWLYLISHMLNPPYGL 60  
 QY 457 FQYSRDDIYTLQINPDSAVNPEHLSYFPRVGRIMGAVFHGHYLDGFTLPEFYQLGKS 516  
 DB 61 FQYSTDNITTLQINPDSINPDHLSYFPRVGRIMGAVFHGHYLDGFTLPEFYQLGKS 120  
 QY 517 ITLDMELVDPDLHNSLVILENDITGVLDHTFCVENHAYGEIIQHEIKPKGKSI PVNEE 576  
 DB 121 IQLSDLESVDPELHKSILVILENDITPVLDHTFCVENHAYGEIIQHEIKPKGKSI PVNEE 180  
 QY 577 NKEEYVRLVYVMRFRGIEAQLQKGFNFVIOHLKTDEKLELITGLCKIDVD 636  
 DB 181 NKEEYVRLVYVMRFRGIEAQLQKGFNFVIOHLKTDEKLELITGLCKIDVD 240  
 QY 637 WKQNTRLXGCTPDSNIVYKFWKAVEFDEERARLLQVTSSSVPLQGFALD---GAA 693  
 DB 241 WKSNTRLXGCTPDSNIVYKFWKAVEFDEERARLLQVTSSTVPLQGFALDGSFGAA 300  
 QY 694 GPRLTTHIQIDACTNLPKATCTNRIDIPYSEYKLYEGLTALIBTCGFAVE 748  
 DB 301 GPRLTTHIQIDACTNLPKATCTNRIDIPYSEYKLYEGLTALIBTCGFAVE 355  
 RESULT 5  
 Q96DE7 ID Q96DE7 PRELIMINARY; PRT; 288 AA.  
 AC Q96DE7;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Similar to E3 ubiquitin ligase SMURF2 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Lung;  
 RA Strausberg R.  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC009527; AA09527.1; -.  
 DR GO: GO:0005622; C:intracellular; IEA.  
 DR GO: GO:0016874; F:ligase activity; IEA.  
 DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO: GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro: IPR000569; HECT domain.  
 DR Pfam: PF00632; HECT; 1.  
 DR SMART: SM00119; HECTC; 1.  
 DR PROSITE: PS00237; HECT; 1.  
 KW LIGase.  
 FT SEQUENCE 1 1

SQ SEQUENCE 288 AA; 33255 MW; FE2B43E300E66537 CRC64;  
Query Match 38.4%; Score 1552; DB 4; Length 288;  
Best Local Similarity 100.0%; Pred. No. 2,4e-113; Indels 0; Gaps 0;  
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 461 RDDIYTLQINPDSAVNPEHLSYFHFVGRIMGAVFHHGHIIDGGFTLPFYKQLLGSIITLD 520  
DB 1 RDDIYTLQINPDSAVNPEHLSYFHFVGRIMGAVFHHGHIIDGGFTLPFYKQLLGSIITLD 60  
QY 521 DMEIYVDPDLNLSVWLTLENDITGVLDHTECVENHNAAGELIOHELIKNGKSIYVNEENKE 580  
DB 61 DMEIYVDPDLNLSVWLTLENDITGVLDHTECVENHNAAGELIOHELIKNGKSIYVNEENKE 120  
QY 581 YVRLVNWFRFLRGIEAOFALQKGFNEVLPQHLKTFDEKELELIIGLGIKIDVNDKVN 640  
DB 121 YVRLVNWFRFLRGIEAOFALQKGFNEVLPQHLKTFDEKELELIIGLGIKIDVNDKVN 180  
QY 641 TRLKCTPDGSIYKMFVKAWEFFDEERRARLQFVTSGRVPLQGFKAQAGAPRLFTI 700  
DB 181 TRLKCTPDGSIYKMFVKAWEFFDEERRARLQFVTSGRVPLQGFKAQAGAPRLFTI 240  
QY 701 HQIDACTNNLPKATCFNRIDIPYSEYKLYEKLTAIEETGFAVE 748  
DB 241 HQIDACTNNLPKATCFNRIDIPYSEYKLYEKLTAIEETGFAVE 288  
RESULT 6  
014326 PRELIMINARY; PRT; 786 AA.  
ID 014326  
AC 014326  
DT 01-JAN-1999 (TrEMBLrel. 09, Created)  
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Ubiquitin-protein ligase.  
GN SPBC169.11c.  
OS Schizosaccharomyces pombe (Fission Yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OC NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972h-;  
RA Volckert G., Wood V., Rajandream M.A., Bartell B.G.;  
RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1-C2 DOMAIN.  
DR EMBL; Z99759; CAB16903.1; -;  
DR PIR; T39585; T39585.  
DR HSP; Q13526; PIN.  
DR GeneDB Sprobe; SPBC169.11c; -;  
DR GO; GO:0005682; C:intracellular; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CALB.  
DR InterPro; IPR000569; HECT\_domain.  
DR InterPro; IPR003469; WW\_domain.  
DR InterPro; IPR001202; WW\_reps\_WWP.  
DR Pfam; PF0168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 3.  
DR PRINTS; P00403; WMDOMAIN.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECT; 1.  
DR SMART; SM00456; WW; 3.  
DR PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS0159; WW\_DOMAIN\_1; 3.  
DR PROSITE; PS0020; WW\_DOMAIN\_2; 3.  
KW Ligase.

SQ SEQUENCE 786 AA; 89259 MW; 57B7A859F5497B9A CRC64;  
Query Match 35.1%; Score 1418; DB 3; Length 786;  
Best Local Similarity 36.7%; Pred. No. 3.2e-102;  
Matches 316; Conservative 127; Mismatches 217; Indels 202; Gaps 20;  
QY 10 GPVRLRLVLCANLVKKDFELDPFPAKYVVGSGGCHSDYTKNTLDPKMHYDLYI 69  
DB 4 GAKVRYVYAADLSKRDLPQDPFALLVDGE-OTHTTKYIKSVNPTWNEGEVTV 62  
QY 70 GKSDSVISVNNHKKIKKQAGAFGCVRLLSNAINELKDTGYORDL-CKLGP 122  
DB 63 KPSSVISIRLFDQKFF-KKQDQFLGLVSF-RMREVSFRSREVSILRPLKSS 115  
QY 123 NDNPTVGRQIVV----- 134  
DB 116 TTNLSVLGNLVKAPSKIRAPAGNHSSTANRTSTPTTTAFTTRTPATNTSN 175  
QY 135 -----SLQSRDRIQGG-----QVWDCSLFDND----- 158  
DB 176 QSTNSTNGTSATSNGTGAGTGAHSRSPYTRQNTGSLSNMHNKMSFEDQY 235  
QY 159 --LPDGEERTASGRIOYLNHTTTTQWERPTRPASEYSPGRPLSCFVDENTPIGCTN 216  
DB 236 GRLPQGERRADSLGRTYVDVHNTRTTW--TRPAS-----STNPVNT- 277  
QY 217 GATCGSSDPRPLARRVRSORRRVMSRTHLTPP-----DLPEGYEQRTQ 263  
DB 278 -----SSD-----SQR-LNHQNR--HLPDSSNPSLMQSDSGNDLPFGEMKRYD 318  
QY 264 QGOVYELHTQGVSTWMDPR-----VPRDLSNINCEELGPPPGWEIR 306  
DB 319 TGRPFVVDHNRTTWDPFRNPVLRPNCGSSVTSGLMQPQLSH-----LGPLSPGEMR 373  
QY 307 NTATGRVYFVDNNHTTFTPTPLRSANHLVLRNQNLKQDQDQVSLCPDDTDECLTVP 366  
DB 374 LTNARVYFVHNNTTTWDDPRLPSAL-----DD-----VP 406  
QY 367 RYKRDLYVKLKLROELSQQCPQAGHCRIEVSRHEIFESYRQVKKRPKDLMKRLMKF 426  
DB 407 QYKCPFRKLIYFRQ-PGMRPLPGQCNVYKRBDIFEDSYAEIIRYSABHLKKRLMIRF 465  
QY 427 RGEGLDYGVARENLYLSHEMLNPYGLFOYSPDDITYTQINDSAVNPEHLSYFHFV 486  
DB 466 DEDDLIDYGLSRSEFFFLSHMPDPICLPEYSADVYTLQINPHSSINSEHNLVFRFI 525  
QY 487 GRIMGAVFHHGHIIDGGFTLPFYKQLLGSIITLDMEIYVDPDLNLSVWLTLENDITGILD 546  
DB 526 GRVIGLAIFFRRFLAFVVSILYKLLKKKYSLADMSIDAEFYRSKLVLENDITGILD 585  
QY 547 HTFCVENHNAAGELIOHELIKNGKSIYVNEENKEVRLVNWFRFLRGIEAOFALQKGFN 606  
DB 586 LTFSEVEDHFGVRIVELITNGENIEVTEENKKYVDLVTEWRVSKVEQGFNAYSGFV 645  
QY 607 EYIPQHLKTFPEKLELIIGLGIKIDVNDKVNRLKCHCPDSIYKMFVKAWEFFDE 666  
DB 646 ELVSPDLVNVFERLELELIGISVDVDEWMSHREYTYIATDVIYKMFETLWGKXE 705  
QY 667 RARLLOFVTSGRVPLQGFKAQAGAPRLFTIHOIDACTNNLPKATCFNRIDIPYE 726  
DB 706 DSKLLQFATGSRIPVNGFRDLQSDGPRKTIKA-GTPQLVATCFRDLDPYV 764  
QY 727 SYEKLYEKLTAIEETGFAVE 748  
DB 765 SKDTLHEKLSLAVENTVPGNE 786  
RESULT 7  
Q8NSA7 PRELIMINARY; PRT; 911 AA.  
ID Q8NSA7  
AC Q8NSA7  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

QY	300	PPGCHIRNATGAGVGVVDHNNRNTQTQPTDPRLSANLHVLRNQNLKQOQOQVYVSLCPDD	359
Db	487	PPGSEERHILHGDGTFYIDHNSKITQWEDPRL-----QN-----	519
QY	360	TECITVPR--YKSDLVQXKLIRBELSOQOPQAGHCRIEVSREEIFEESYQVAKPKD	417
Db	520	-PAITGAVAYSNRPFQKXDYPRKKLKKRPADIPIRFEKKLHRRNIPEESFRRLMSVRPD	578
QY	418	LWK-RLMIKRGEEGIDYGVAREVLVYLISHMLNPYGLGFQYSRDDIYTLQINPDSAV-	475
Db	579	VLFARMIIEESKSGIDYGVARERWFLFSKEMRPYVYGLFEYSATNNTYLQINPNSGLC	638
QY	476	NPEHLSYFHVGRIMGAVFHCHYIDGQFTLPFKYOLGKSIITDDMELVDPRLHSLVW	535
Db	639	NEDHLSYFTTIGVAGSLAFVHRKLLDGFIRPFKMLGQITLINDHESVDSYNSLKM	698
QY	536	ILENDITGVLDHTPCVENHAYGEIIOHELKPNKGSIPVNEENKEYYRLVYNNRPLRGIE	595
Db	699	ILENDPT-ELIDLMFCIDENFEGQTQVDLKPNGSEIMWTNENKREYIDLVIQWRFYNRVQ	757
QY	596	AGFLAQKGNGNEVPIPHILKTFDEKEXELIIIGLGIKIDVNDWKYNRLLK--HCTPDNSNTV	653
Db	758	KQNNALPBEGETELPDLIKTFDEBELMLCGLADVDVNDKROHSIYKNGYV-PVHPVYI	816
QY	654	KMTFWAVEFDEERRARLIQFYTGSSRRVPLQGFALQGAAPRLFTIHQIDACTNNLPKA	713
Db	817	QWFMKAVLLMDAEKRIRLIQFYTGSSRRVPMGFAELVGSNQPQFTIEQVGS-PEKLPR	875
QY	714	HTCFNRIDIPPESEYKLYEKLLTIESTCF 745	
Db	876	HTCFNRIDIPPEYETFEDELREKLMAVENAQGF 907	

RESULT 8

Q96PUS PRELIMINARY; PRT; 955 AA.

AC Q96PUS 01-DEC-2001 (TEMBLrel. 19, Created)

DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)

DE NEDD4-like ubiquitin ligase 3.

GN NEDL3

OS Homo sapiens (human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Okamoto Y., Miyazaki K., Sakamoto M., Kato C., Nakagawara A.;

RT "homo sapiens NEDD4-like ubiquitin ligase 3.";

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=21269431; PubMed=11244092;

RA Harvey F.K., Dindom A., Cook I.D., Kumar S.;

RT "the Nedda4-like protein KIAA0439 is a potential regulator of the epithelial sodium channel.";

RL J. Biol. Chem. 276:8597-8601 (2001).

RN [3]

RP SEQUENCE FROM N.A.

RA MEDLINE=20501262; PubMed=11046148;

RA Winberg G., Matskova L., Chen F., Plant P., Rotin D., Gish G.,

RA Ingbar R., Erbergh I., Pawson T.,

RT "latent membrane protein 2A of Epstein-Barr virus binds WW domain E3 protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases.";

RL Mol. Cell. Biol. 20:8526-8535 (2000).

CC -1- SIMILARITY: CONTAINING 1 C2 DOMAIN.

DR EMBL: AB071179; BAB69424.1;

DR GO: GO:0005622; C:intracellular; IEA.

DR GO: GO:0016874; F:ligase activity; IEA.

DR GO: GO:0004842; F:ubiquitin-protein ligase activity; IEA.

DR GO: GO:000512; F:ubiquitin cycle; IEA.

DR InterPro: IPR000008; C2.

```

DR InterPro; IPR008973; C2_CaB.
DR InterPro; IPR00569; HECT_domain.
DR InterPro; IPR002349; WW.
DR InterPro; IPR01202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2_1.
DR Pfam; PF00632; HECT_1.
DR Pfam; PF00397; WW_4.
DR PRINTS; PR00360; C2DOMAIN.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2_1.
DR SMART; SM00119; HECTC_1.
DR SMART; SM00456; WW_4.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS02337; HECT_1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS00020; WW_DOMAIN_2; 4.
DR LIGASE.
KW LIGASE.
SQ SEQUENCE 955 AA; 110021 MW; A8BB278A37F6A6B5 CRC64;

Query Match 34.3%; Score 1386; DB 4; Length 955;
Best Local Similarity 34.4%; Pred. No. 1,4e-99;
Matches 330; Conservative 125; Mismatches 247; Indels 258; Gaps 26;

14 LRLVLCANLVKCDFFRLPDPFAKV---VVDGSGGCH--STDVYKNTLDPKKNQHDLY 68
22 LRVKVVSGIDLAKDIFGASDPYKLSLYVDENRELALVQTKTKTLNPKMNEFFYR 81
69 IGKSD-SVTISVNNHKKHKKGAGFLGCVR-----LISNAIN 105
82 VNPSNHRLLFEVVDENRLTRDD---FLGQVDVPLSHLPTEDPTMERPYTFKDFLLRPSH 138
106 RLKDTGYORLDCKLGPNDNDTVRGQIVVSLQSRDRIIGTGQVDCSRLFPND----- 158
139 KSRVKGFLRLKMAVYFKNG-----GODENSDQRDMEHGEVVD-----SNDASQHOE 188
159 -----LPGMEERRTASGRIOYLNHITRTQWERPT----- 189
189 ELPPPLPFGMEBEKVDNLGRITYVNNHNRRTQWHRPSLMDVSSDNNIRQINQEAHRR 248
190 ----RPASEYSSP-----GRPLSCFVDENPTIGTN-----GATCGQSSDPR-- 227
249 FRSRHHISEDLPEPSEGGDVPEPWEPT-ISEEVNIAGSLGLALPPPPASGSRTPQEL 307
228 -----LAER-----RVRS-----QRHNNYSRTHLHTPPDL 253
308 SEELSRRLQITPPDSNGEQFSSLIQREBSSRLRSCSVTDAVAEQGHLPPPSVAVHTTPEL 367
254 PEGYEQRTTQGGQVYFLHTQGVSTWHD-----RVPRDL 288
368 PFGMEERDAKGRITYVNNHNRRTTWTPTMQLAEDGASGATSNNNHLIEPQIRRPRL 427
289 SNINCEELGP-----GODENSDQRDMEHGEVVD-----SNDASQHOE 188
428 SSPVTLSAPLEGAKDSPVRAVKDTLSNPQSPQSPYNSPKPOHKYTGQSLPPGMEWRI 487
308 TATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQLKQOQOQVSLCPDD----- 359
488 APNGRPFIDHNTKTWTWEDPRLLKFPVHM-----RSKSTLNPNDLGPRLPPGM 534
360 -----TECLTVPR--YKRDVQLKTLRQLSLQOQOQ 389
535 EERHLDGRFFYVDHNSKITQWEDPRLQNPATIGPAPVSRERKQKYDVFRRKLKAPAI 594
390 AGHCRLEVSREELFEESYQVMKRPKDLWK-RLMTKFRGEEGLDYGVAREWLYLLSH 448
595 PNFPEKHLRNHNFEESYRIMSVPKPDVAKALWIEFESEKGLDYGVAREWFLSK 654
449 MLNPYGLFQYSRDIITLQINPDSAV-NPEHLSTPHFVGRINGMVFCHYIDGGFTLP 507
655 MFNPYGLFESATDNTLQINPNSGLCNEDHLSYFTFICRVAGLVPFGKLLDGGFIR 714
508 FYKOLGKSTILDMDLVDLPHNSLWLIENDITGVLDHTPCVENNAVGEIIQHEIKDN 567

```

```

DB 715 FYKMLGKQITLNDMESVDSEFYNSLKNITLNDPT-ELDMFCIDENPFQATQVMDIKN 773
QY 568 GKSIPVNEENKKEVYRLVYNNRFLRGIEAQLAQGFNEVYIPOHLKTFDEKELEIIC 627
DB 774 GSEIMVTNNKREYIDLVIQWRFVNRVQKQNNAFLEGFTELLPDLIKIFDENLELIMC 833
QY 628 GLGKIDVNDKMNTRLK--HCTPDSNIVKWEKVAVEFPDEERRARLLQFTGSSRFVLOG 685
DB 834 GLGSDVDVNDWQSHIYNQGYC-PNHPIVQWFKAVLLMDAKRIRLQFTGSRVPNG 892
QY 686 FKALQGAAPRLFTIHQIDACTNNLPKATCFNRRIDIPYSEYKLYEKLITALETCGF 745
DB 893 FAELYGNSGQLFTIEQWGS-PEKLPRAHICFNRLLDPPETPEDREKILMAVENAQGF 951

RESULT 9
Q725F1 PRELIMINARY; FR; 947 AA.
AC Q725F1
DT 01-OCT-2003 (TrEMBLrel. 25. Created)
DT 01-OCT-2003 (TrEMBLrel. 25. Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25. Last annotation update)
DE Ubligulin ligase NEDD4L.
GN NEDD4L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxId=9606;
RN (1)
RP SEQUENCE FROM N.A.
RA Qi H., Labrie C.;
RT "New splicing isoform of human Ned4-2."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY12985; AAM76730.1; -.
KW LIGASE.
SQ SEQUENCE 947 AA; 109429 MW; 95CB0FDEDCB96639 CRC64;

Query Match 34.3%; Score 1384; DB 4; Length 947;
Best Local Similarity 34.4%; Pred. No. 2e-99;
Matches 330; Conservative 125; Mismatches 247; Indels 258; Gaps 26;

14 LRLVLCANLVKCDFFRLPDPFAKV---VVDGSGGCH--STDVYKNTLDPKKNQHDLY 68
14 LRVKVVSGIDLAKDIFGASDPYKLSLYVDENRELALVQTKTKTLNPKMNEFFYR 73
69 IGKSD-SVTISVNNHKKHKKGAGFLGCVR-----LISNAIN 105
74 VNPSNHRLLFEVVDENRLTRDD---FLGQVDVPLSHLPTEDPTMERPYTFKDFLLRPSH 130
106 RLKDTGYORLDCKLGPNDNDTVRGQIVVSLQSRDRIIGTGQVDCSRLFPND----- 158
131 KSRVKGFLRLKMAVYFKNG-----GODENSDQRDMEHGEVVD-----SNDASQHOE 180
159 -----LPGMEERRTASGRIOYLNHITRTQWERPT----- 189
181 ELPPPLPFGMEBEKVDNLGRITYVNNHNRRTQWHRPSLMDVSSDNNIRQINQEAHRR 240
190 ----RPASEYSSP-----GRPLSCFVDENPTIGTN-----GATCGQSSDPR-- 227
241 FRSRHHISEDLPEPSEGGDVPEPWEPT-ISEEVNIAGSLGLALPPPPASGSRTPQEL 299
228 -----LAER-----RVRS-----QRHNNYSRTHLHTPPDL 253
300 SEELSRRLQITPPDSNGEQFSSLIQREBSSRLRSCSVTDAVAEQGHLPPPSVAVHTTPEL 359
254 PEGYEQRTTQGGQVYFLHTQGVSTWHD-----RVPRDL 288
368 PFGMEERDAKGRITYVNNHNRRTTWTPTMQLAEDGASGATSNNNHLIEPQIRRPRL 419
289 SNINCEELGP-----GODENSDQRDMEHGEVVD-----SNDASQHOE 188
420 SSPVTLSAPLEGAKDSPVRAVKDTLSNPQSPQSPYNSPKPOHKYTGQSLPPGMEWRI 479

```

```

Qy 308 TATGTVFVVDHNNHTTQFTDPRLSANLHLVINRONQLKDOQQQVSLCPDD----- 359
Db 480 APNGRPFPEIDHNTKTITMEDPRLKFPVHM-----SKSLINPNDLGPPLPGW 526
Qy 360 -----TECLTVPR--YKRDLVQKXKILROELSOQOPQ 389
Db 527 EERIHLDGRTYIDHNSKITQWEDPRLONPAITGPAPYSGREFQKIDYRKKLKKRADI 586
Qy 390 AGHCRIEVSREIEESYRQVMKRPYOLMK-RIMIKRGEGLDYGVAREMILYLSHE 448
Db 587 PNREMKLHRNNIEEESYRIRMSYKRPDYKARIMIEESEKIGDYGVAREMFLISKE 646
Qy 449 MLNRYGGLFOYSRDITLQINPDASV-NPEHLSYFHVGRIMGMAVFEHGYIDGFTLP 507
Db 647 MFNPYIGLFEYSADNTLQINPNSGLCNEDHLSYFTFGVAGLAFVHFKLLDGFPIRP 706
Qy 508 FYKOLGKSLITLDMELVDPDLHNSLWVILENDITGVLDHTFCVENHAYGEIIOHEIKPN 567
Db 707 FYKMLGKQITLNDMESVDSYNSLKWILENDPT-ELDLMFICIDENFGQYQVLDKPN 765
Qy 568 GKSIPVNEKKEVYRKYVMRFLRGIEAQLAKQKFNVEVIRPHLKTDEKLELITC 627
Db 766 GSEIMVTNENREYIDVIOVRFYNRVQKQNNALFEGFTLLPDLIKIDENELMLMC 825
Qy 628 GLGKIDVDMKVNTRLK-HCTPDSNIVKMWKXAVEFDEERARLLQFTGSSRPVLOG 685
Db 826 GLGVVDVDMRGHILYKNGYC-PHNPVIOFWKAVLMDLMDKRLILQFTGSRVPMNG 884
Qy 686 FKALQGAAPFLFTIHQIDACTNNLPKATCFNRIDIPYSEYEKYLEKLLTIEETCGF 745
Db 885 FAEIYGSNGPOLFTIEQWGS-PEKLPRAHFCFNRILDLPYETEDLREKLMAYENAGOF 943

```

## RESULT 10

```

Q8ND8      PRELIMINARY;      PRT;      295 AA.
AC Q8ND8;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP564H23.3
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP TISSUE=Brain;
RC TISSUE=Brain;
RA Wanduc R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL84242; CAD38919.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0005512; P:ubiquitin cycle; IEA.
DR InterPro; IPR00569; HECT_domain.
DR Pfam; PF00632; HECT_1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS50237; HECT; 1.
KW Hypothetical protein.
FT NON TER
FT 1
SQ SEQUENCE 295 AA; 34030 MW; 0507325127A943EA CRC64;

```

```

Query Match 34.2%; Score 1379; DB 4; Length 295;
Best Local Similarity 85.4%; Pred. No. 9e-100;
Matches 251; Conservative 24; Mismatches 19; Indels 0; Gaps 0;

```

```

Qy 455 GLFOYSDDIYTLQINPDASVNPBHL-SYFHVGRIMGMAVFEHGYIDGFTLPFYKOLG 514
Db 2 GLPQYSTDNIMQINPDSSINPDHLSYFHVGRIMGMAVFEHGYIDGFTLPFYKOLG 61
Qy 515 KSITLDMELVDPDLHNSLWVILENDITGVLDHTFCVENHAYGEIIOHEIKNGKSI PVN 574

```

```

Db 62 KPIQLSDLESVDPELHKLWILENDITPVLDHTFCVENHAFGLILOHEIKPNRNPVPT 121
Qy 575 EENKKEYYRLYVMWFFLNGIEAQLAKQKFNVEVIRPHLKTPEKELELITCGAGKIDV 634
Db 122 EENKKEYYRLYVMWFFLNGIEAQLAKQKFNVEVIRPHLKTPEKELELITCGAGKIDV 181
Qy 635 NDWKNTLKLKCTPDSNIVKMWKXAVEFDEERARLLQFTGSSRPVLOGFKALQGAAG 694
Db 182 NDWKNTLKLKCTPDSNIVKMWKXAVEFDEERARLLQFTGSSRPVLOGFKALQGAAG 241
Qy 695 PRLFTIHQIDACTNNLPKATCFNRIDIPYSEYEKYLEKLLTIEETCGAVE 748
Db 242 PRLFTIHQIDACTNNLPKATCFNRIDIPYSEYEKYLEKLLTIEETCGAVE 295

```

## RESULT 11

```

Q9CSE3      PRELIMINARY;      PRT;      258 AA.
AC Q9CSE3;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE 2810411E22Rik protein (Fragment).
GN 2810411E22Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.;
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RA MEDLINE=5781/6J; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojopori T., Sono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gasteierland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,
RA Hayashizaki Y.;
RL "Functional annotation of a full-length mouse cDNA collection.";
RT Nature 409:685-690(2001).
DR EMBL; AK013082; BAB28637.1; -.
DR MGI; MGI:1913563; 2810411E22Rik.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0005512; P:ubiquitin cycle; IEA.
DR InterPro; IPR00569; HECT_domain.
DR Pfam; PF00632; HECT_1.
DR SMART; SM00119; HECTC; 1.
DR PROSITE; PS50237; HECT; 1.
FT NON TER
FT 1
SQ SEQUENCE 258 AA; 29670 MW; B0B804BC1FCB98 CRC64;

```

```

Query Match 34.1%; Score 1378; DB 11; Length 258;
Best Local Similarity 98.8%; Pred. No. 8.9e-100;
Matches 255; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 491 GMAVPHGYITGGTLPFYKOLGKSTLDDMEVYDPDLHNSLWVILENDITGVLDHTFC 550
Db 1 GMAVPHGYITGGTLPFYKOLGKSTLDDMEVYDPDLHNSLWVILENDITGVLDHTFC 60

```

```

QY 551 VEHNAVGEIIQHLEKNGKSIPIVNEENKKEVYRLVYNNMFLRGIEAOFALQKGFNEYIP 610
DB 61 VEHNAVGEIIQHLEKNGKSIPIVNEENKKEVYRLVYNNMFLRGIEAOFALQKGFNEYIP 120
QY 611 QHLKTFDEKELELLICGLKIDVANDMKNTLKTCTPDSNVYKTFWKAVERFDEERAR 670
DB 121 QHLKTFDEKELELLICGLKIDVANDMKNTLKTCTPDSNVYKTFWKAVERFDEERAR 180
QY 671 LLOFVTSGRVPLQGFKAQGAAPRLFTIHQIDAQTNLPRKHTCFNRIDIPYSEYK 730
DB 181 LLOFVTSGRVPLQGFKAQGAAPRLFTIHQIDAQTNLPRKHTCFNRIDIPYSEYK 240
QY 731 LYEKLTALIEETCGFAVE 748
DB 241 LYEKLTALIEETCGFAVE 258

RESULT 12
Q725N3 PRELIMINARY; PRT; 975 AA.
ID Q725N3;
AC Q725N3;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE NEDD4.2
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Maibert-Colas L., Nicolas G., Galand C., Lecomte M.-C., Dhery D.;
RT "Identification of new partners of the epithelial sodium channel alpha
RT subunit."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY312514; AAP75706.1; -.
SQ SEQUENCE 975 AA; 111930 MW; AFED52AD504587B8 CRC64;

Query Match 34.0%; Score 1371; DB 4; Length 975;
Best Local Similarity 33.3%; Pred. No. 2, 1e-98;
Matches 332; Conservative 121; Mismatches 233; Indels 310; Gaps 26;

QY 14 LRLTVLCAKLVKDDFFRLPDPFAKV--VVDGSGQCH--STDVYKNTLDPKMNQHYDLY 68
DB 22 LRVKVVSGIDLAKDIFGASDPYVKLSIYADENRRLALVQTKIKTLNPKMNEEFYFR 81
QY 69 IGKSD-SVTISVNNHKKIKHKGAGFLGCVR-----LUSNAIN 105
DB 82 VNPFSNHLLEFVFDENRRLTRDD--FLGQVDPVLSHLPEDPTMERPRYTFKDFLLRPRSH 138
QY 106 RLKDTGYQRDLCKLGNNDNTVAGQIVASLQSRDRIQGGVAVDQSRLEFND----- 158
DB 133 KSRVKGFLRLKAMVPPNG-----GQDENSDQORDMEHGEVVD-----SNDASQHQK 188
QY 159 -----LPGWBERRTASGRIOYLNIHTTTQWERPT----- 189
DB 189 ELPPPLPPLPGMEERKVDNLGRITVYNNHNRRTQWHRPSLMDVSSDNNIRQINQEAHR 248
QY 190 -----RPASE-----YSPG----- 199
DB 249 FRGRHRISLELPEPBGSGDVPEPEMETISEVINIAGDSLGLALPPPPASGKTSQELIS 308
QY 200 -----RPLSCFV----- 206
DB 309 EELSRRLQITPDSNGEQFSSLIQREPSRLRSCSVTDVAEQLLPPPSAPAGRASSTY 368
QY 207 -----DENP-----ISGTMGATCG----- 221
DB 369 TGGEEPTPSVAYVHTTGLPSGWEERKDAKGRITVYNNHNRRTTWRPIQLAEDGASG 428
QY 222 -QSGDPLAERVRVQSHRYMSRT----- 245
DB 429 ATNSNNHLIPQIR--RPRSLSPFTVLSAPLEGAKDSFVRBAVKDTLSNPSQPSPTN 486

```

```

QY 246 -----HLHTPPDLPEGEYORITQOGQVYFLHQTGVSTWHDPRVP-----RDISNINCEE 295
DB 487 SPKPKHKTQSFLPPLPGWEMRIAPNGRPFFIDHNTKTTTWEDPRLKFPVHRKSTSLNPN 546
QY 296 LGRPLPGMEIRNATGTVYFVDHNNRTTQFTDPLSANHLVLVNRQQLNDQOQGVSL 355
DB 547 LGRPLPGMEIRHLDGTFYIDHNSKITQWEDPRL-----QV----- 583
QY 356 CPDDTECLTYPR--YKRDLYQKLIHQELSQOQPAQGHCRIEYSREIFEESRYQWKM 413
DB 584 -----PAITGPAVAYSREFQKDYPERKCLKKPADINRFEMKLRHNNIFEESYRIMSV 638
QY 414 RPKDLMK-RLMKRGRGEGLDYGVANEMTYLISHENLNTYTGIFQYSRDDITLQINPD 472
DB 639 KRDPVLKARLMIPESESKGLDYGVAAREWFFLSKEMFNRYGLFEYSATDNTYLLQINPN 698
QY 473 SAV-NPEHLSYFHFVGRIMGKAVPHGHYIDGFTLPPFYKOLLGKSIITLDMELVDPDLN 531
DB 699 SGLCNEDHLSYFTTIGVAGIAPHGLDQFFIRPYKMLKSKQITLNDMESVDSRYN 758
QY 532 SLVWILENDITGVLDHTFCVEHNAVGEIIQHEKPNKKSIPVNEENKKEVYRLVYNNRFL 591
DB 759 SLKWLLENDPT-ELDLMFCIDEENFGQTYQVDLKFNGSEIMVTNENKREYIDVLTQWRFV 817
QY 592 RGIQAQFLALQKGFNEYIPQHLKTFDEKELELLICGLKIDVANDMKNTLKTCTPDSNVYKTFWKAVERFDEERAR 649
DB 818 NRVQKQMAFLGEGTELLPDLIKIPDENBELMCGLDVANDMKNTLKTCTPDSNVYKTFWKAVERFDEERAR 876
QY 650 SNIVKMFKAVERFDEERARLLOFVTSGRVPLQGFKAQGAAPRLFTIHQIDAQTN 709
DB 877 HPVIOFWKAVLMDAERIRLLQFTVTSRVPNGFAELYGNGPQLFTIEQWGS-PEK 935
QY 710 LPKAHTCFNRIDIPYSEYKLYEKLTALIEETCGF 745
DB 936 LPKAHTCFNRIDIPYSEYKLYEKLTALIEETCGF 971

RESULT 13
Q725F2 PRELIMINARY; PRT; 967 AA.
ID Q725F2;
AC Q725F2;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ubiquitin ligase NEDD4G.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Qi H., Labrie C.;
RT "New splicing isoform of human Nedd4-2."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY112984; AAM76729.1; -.
KW LIGase.
SQ SEQUENCE 967 AA; 111339 MW; 86940A75880539F7 CRC64;

Query Match 33.9%; Score 1369; DB 4; Length 967;
Best Local Similarity 33.3%; Pred. No. 3e-98;
Matches 332; Conservative 121; Mismatches 233; Indels 310; Gaps 26;

QY 14 LRLTVLCAKLVKDDFFRLPDPFAKV--VVDGSGQCH--STDVYKNTLDPKMNQHYDLY 68
DB 14 LRVKVVSGIDLAKDIFGASDPYVKLSIYADENRRLALVQTKIKTLNPKMNEEFYFR 73
QY 69 IGKSD-SVTISVNNHKKIKHKGAGFLGCVR-----LUSNAIN 105
DB 74 VNPFSNHLLEFVFDENRRLTRDD--FLGQVDPVLSHLPEDPTMERPRYTFKDFLLRPRSH 130
QY 106 RLKDTGYQRDLCKLGNNDNTVAGQIVASLQSRDRIQGGVAVDQSRLEFND----- 158

```

```

Db 131 KSRVKGFLRLKMAVYKPKG-----GQDENSDQDDMEHGWVD-----SNDASQHOE 180
Qy 159 -----LPDGEERTASGRIOYLNHTRTTTOWMEPT----- 189
Db 181 ELPPPLPPGWEKVDNIGRTYYVNNHNRRTTQWHRPSIMDVSSDNNIRQINQEAHRR 240
Qy 190 -----RPASE-----YSSPG----- 199
Db 241 FRSRHISEDLEPPESSEGVDPEEWETISEEWINAGSLGLALPPPPASGSRTPQELS 300
Qy 200 -----RPLSCFV----- 206
Db 301 EELSRRLQITPDSNGCFSSLIQREPSRSLSSCVTDVAVAQCHLPPPSAPAGRASSTV 360
Qy 207 -----DENTP-----IGTNGATG----- 221
Db 361 TGESEPTSAVAYHTTTPGLPSGWEERKADKGRYYVNNHNRRTTTPPIVQLAEDGASGS 420
Qy 222 -QSSDPRLAERRVNSQRHRYMSRT----- 245
Db 421 ATNSNNHLIEPQIR--RPRSLSPVTLASPLLEGAKDSPVRAVYKDTLSNPQSPQEPYN 478
Qy 246 -----HLTPPDLPEGIEQRTTQOGQYFHTQTGVSTWMDPRVP-----RDLSNINCEE 295
Db 479 SPKQHKYKTQGLPPGWEEMRIAPNGRPFIDHNTKTTWEDPRLKFPVHNRKSTLSLND 538
Qy 296 LGPLPGEWEIRNTATGRVYFVDHNNRTTQFTDPLSLANLHLVNRQQLDQOQOQVSVL 355
Db 539 LGPLPGEWEIRIHDGRTFYIDHNSKITQWEDPL-----QV----- 575
Qy 356 CPDTECTVPR--YKRDLYOKLILQELSQOQPOAGHCRIEVSREEIFEESYQVYMK 413
Db 576 -----PATGPVAVYSSEFKQYDFPKKPKKPADINREEMKLRNNIPEESYRRIMSV 630
Qy 414 RPKDLWK-RLMIKRSGEGLDYGVAREWLYLSHEMLNPPYGLFOYSRDDIYTLQINPD 472
Db 631 KRPDLKRLWIESEKGLDYGVAREWFLSKENFPYGLFEISATDNTTLQINPN 600
Qy 473 SAV-NPEHLSYFHFVGRIMGMAVPHGYIDGFTLLPYKQLIGKSTLDDMELVDDPLN 531
Db 691 SGLCNEDHLSYFTIGRVAAGLAVPHGKLDGFFIRPYKMLGKQITLNDMESVDSBYN 750
Qy 532 SLWMLNDITGVLDHIFCVENHAYGELIQELKPKNKSTIPVBNENKEVRYLYVMRPL 591
Db 751 SLKMLILENDPT-ELDLMFCIDEENFGQYQVDLKNPSEIMVTNENKREYIDLVIQMRFV 809
Qy 592 RGIAPQFLALQKGENEYIPOHILKTPDEKELELIIIGLGIKIDVNDWKVNTRLK--HCTPD 649
Db 810 NRVCQKQMAFLFEGTELLPDLIKIPDENLELIMGLGDVDNDRQSHIYNQYIC-PV 868
Qy 650 SNIVKWKVAVFEDDERARALLOFVYSSRVPLQGFKALQGAAGPRLFTIHQIDACTNN 709
Db 869 HPVIOQWPKAVLMDAKIRILLOFVGTGRVNPNGPAELVGSNGPOLFTIEOWGS-PEK 927
Qy 710 LPKAHTCFNRIDIPYSEYEKYLTLTAIEETCGF 745
Db 928 LPRAHCTCFNRDLPPYETFEELRKLLMAVENAGF 963

```

## RESULT 14

```

Q9NT88 PRELIMINARY; PRT; 820 AA.
ID Q9NT88;
AC Q9NT88;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKRZP434P2422.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Blum H., Baueersachs S., Wewes H.W., Gassenhuber J., Wiemann S.,
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
DR EMBL; AL137469; CAB70754.1; -.
DR PIR; T46412; T46412.
DR HSSP; Q33526; 1PIN.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaB.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR002349; WW_Rsp5_WWP.
DR Pfam; PF00168; C2_1.
DR Pfam; PF00632; HECT_1.
DR Pfam; PF00397; WW_3.
DR SMART; SM00403; WMDOMAIN.
DR SMART; SM00239; C2_1.
DR SMART; SM00115; HECTC_1.
DR SMART; SM00436; WW_3.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT_1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS50020; WW_DOMAIN_2; 3.
KW Hypothetical protein.
FT NON_TER
SQ
SEQUENCE 820 AA; 95283 MW; 0FDB34B29B3F4123 CRC64;

Query Match 33.8%; Score 1364.5; DB 4; Length 820;
Best Local Similarity 37.1%; Pred. No. 5,4e-98;
Matches 315; Conservative 115; Mismatches 221; Indels 199; Gaps 25;

Qy 50 TDVYNTLDPKRNQHYDLYIKSD-SVATSYVNNKHKKGAQFLGVR----- 98
Db 12 TKTKIKTLNPKWNEEFYRVNPSNHRLLFVFDENRLTRD--FLGVDVPLSHLPED 68
Qy 99 -----LISNAINRLKDTGYQRLDCKLGPNDNDTVRQIVVSLQSRDRTG 146
Db 69 PTMERPTFFKDFLPRSHKSRVAGFLRKMAVYKPKG-----QDENSDQDDMEHGW 123
Qy 147 QVYDCSLFND-----LPDGEERTASGRIOYLNHTRTTTOWMEPT----- 189
Db 124 EVVD-----SNDASQHOEBLPPPLPPGWEKVDNIGRTYYVNNHNRRTTQWHRPSIMDV 178
Qy 190 -----GATCGQSSDPRLAE--RVR-----SGRHNYSRTHL----- 216
Db 179 SSESNNIRQINQEAHRRFRSRHISEDLEPPESSEGVDPEEWET-ISEEWINAGSISLG 237
Qy 217 -----GATCGQSSDPRLAE--RVR-----SGRHNYSRTHL----- 247
Db 238 LALPPPPASPGSRTPSDELSEELSRRLQITPDSNGEQFSSLIQREPSRSLSSCVTDVA 297
Qy 248 -----HTPPD----- 266
Db 298 EOGHLPPPAKDSPYRAVYKDTLSNPQSPQEPYNSPKQHKVTQSFAPPGWEMRIAPNGR 357
Qy 267 VYFHTQTGVSTWMDPRVP-----RDLSNINCEEAGLPPGWEIRNTATGAYVYFVDHNR 321
Db 358 PFFIDHNTKTTWEDPRLKFPVHNRKSTLSLNPNDLGLPPGWEIRIHDGRTFYIDHNSK 417
Qy 322 TTQFTDPLSLANLHLVNRQQLDQOQOQVSVLCPDTECTVPR--YKRDLYOKLIL 379
Db 418 ITQWEDPL-----QV----- 449
Qy 380 ROELSQOQPOAGHCRIEVSREEIFEESYQVYMKRPRDLWK-RLMIKRSGEGLDYGVVA 438
Db 450 RKKLKKPADINREEMKLRNNIPEESYRRIMSVYKRDVILKARIMIEFESKGLDYGVA 509
Qy 439 RWMYLLSHEMLNPPYGLFOYSRDDIYTLQINPDNAV-NPEHLSYFHFVGRIMGMAVPHG 497

```



*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:40:45; Search time 23.1006 Seconds  
(without alignments)  
1615.783 Million cell updates/sec

Title: US-10-009-945-2  
Perfect score: 3884  
Sequence: 1 GGGSSIKRLTVCAKMLAKK.....EKLYEKLTAVBETGFAVE 723

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

1: Issued Patents AA:\*  
2: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/6C\_COMB.pep:\*  
7: /cgn2\_6/prodata/2/1aa/6D\_COMB.pep:\*  
8: /cgn2\_6/prodata/2/1aa/6E\_COMB.pep:\*  
9: /cgn2\_6/prodata/2/1aa/6F\_COMB.pep:\*  
10: /cgn2\_6/prodata/2/1aa/6G\_COMB.pep:\*  
11: /cgn2\_6/prodata/2/1aa/6H\_COMB.pep:\*  
12: /cgn2\_6/prodata/2/1aa/6I\_COMB.pep:\*  
13: /cgn2\_6/prodata/2/1aa/6J\_COMB.pep:\*  
14: /cgn2\_6/prodata/2/1aa/6K\_COMB.pep:\*  
15: /cgn2\_6/prodata/2/1aa/6L\_COMB.pep:\*  
16: /cgn2\_6/prodata/2/1aa/6M\_COMB.pep:\*  
17: /cgn2\_6/prodata/2/1aa/6N\_COMB.pep:\*  
18: /cgn2\_6/prodata/2/1aa/6O\_COMB.pep:\*  
19: /cgn2\_6/prodata/2/1aa/6P\_COMB.pep:\*  
20: /cgn2\_6/prodata/2/1aa/6Q\_COMB.pep:\*  
21: /cgn2\_6/prodata/2/1aa/6R\_COMB.pep:\*  
22: /cgn2\_6/prodata/2/1aa/6S\_COMB.pep:\*  
23: /cgn2\_6/prodata/2/1aa/6T\_COMB.pep:\*  
24: /cgn2\_6/prodata/2/1aa/6U\_COMB.pep:\*  
25: /cgn2\_6/prodata/2/1aa/6V\_COMB.pep:\*  
26: /cgn2\_6/prodata/2/1aa/6W\_COMB.pep:\*  
27: /cgn2\_6/prodata/2/1aa/6X\_COMB.pep:\*  
28: /cgn2\_6/prodata/2/1aa/6Y\_COMB.pep:\*  
29: /cgn2\_6/prodata/2/1aa/6Z\_COMB.pep:\*  
30: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
31: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
32: /cgn2\_6/prodata/2/1aa/6C\_COMB.pep:\*  
33: /cgn2\_6/prodata/2/1aa/6D\_COMB.pep:\*  
34: /cgn2\_6/prodata/2/1aa/6E\_COMB.pep:\*  
35: /cgn2\_6/prodata/2/1aa/6F\_COMB.pep:\*  
36: /cgn2\_6/prodata/2/1aa/6G\_COMB.pep:\*  
37: /cgn2\_6/prodata/2/1aa/6H\_COMB.pep:\*  
38: /cgn2\_6/prodata/2/1aa/6I\_COMB.pep:\*  
39: /cgn2\_6/prodata/2/1aa/6J\_COMB.pep:\*  
40: /cgn2\_6/prodata/2/1aa/6K\_COMB.pep:\*  
41: /cgn2\_6/prodata/2/1aa/6L\_COMB.pep:\*  
42: /cgn2\_6/prodata/2/1aa/6M\_COMB.pep:\*  
43: /cgn2\_6/prodata/2/1aa/6N\_COMB.pep:\*  
44: /cgn2\_6/prodata/2/1aa/6O\_COMB.pep:\*  
45: /cgn2\_6/prodata/2/1aa/6P\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2913.5	75.0	735	US-08-539-205A-2	Sequence 2, Appl
2	2913.5	75.0	735	US-08-392-163A-2	Sequence 2, Appl
3	1471	37.9	766	US-08-539-205A-4	Sequence 4, Appl
4	1471	37.9	766	US-08-392-163A-4	Sequence 4, Appl
5	1318	33.9	854	US-09-070-060-4	Sequence 4, Appl
6	1318	33.9	854	US-09-357-746-4	Sequence 4, Appl
7	1301	33.5	852	US-09-070-060-3	Sequence 3, Appl
8	1301	33.5	852	US-09-357-746-3	Sequence 3, Appl
9	1259	32.4	927	US-08-895-601-6	Sequence 6, Appl
10	1228.5	31.6	834	US-08-539-205A-6	Sequence 6, Appl
11	1228.5	31.6	834	US-08-392-163A-6	Sequence 6, Appl
12	1225	31.5	906	US-08-630-916A-48	Sequence 48, Appl
13	1047	27.0	683	US-08-630-916A-46	Sequence 46, Appl
14	530.5	13.7	874	US-08-247-904B-8	Sequence 8, Appl
15	530.5	13.7	874	US-08-767-942A-21	Sequence 21, Appl
16	509	13.1	866	US-08-100-692-1	Sequence 1, Appl
17	509	13.1	866	US-08-674-030-1	Sequence 1, Appl
18	473	12.2	1083	US-08-895-601-5	Sequence 5, Appl
19	241	6.2	486	US-08-348-518C-2	Sequence 2, Appl
20	238.5	6.1	472	US-08-348-518C-5	Sequence 5, Appl
21	238.5	6.1	472	US-08-476-509B-5	Sequence 5, Appl
22	162.5	4.2	54	US-08-630-916A-124	Sequence 124, App
23	162	4.2	448	US-08-476-509B-2	Sequence 2, Appl
24	161.5	4.2	454	US-08-348-518C-4	Sequence 4, Appl
25	161.5	4.2	454	US-08-476-509B-4	Sequence 4, Appl
26	154.5	4.0	55	US-08-630-916A-75	Sequence 75, Appl
27	148.5	3.8	58	US-08-630-916A-84	Sequence 84, Appl

28	148	3.8	51	US-08-630-916A-117	Sequence 117, App
29	147.5	3.8	54	US-08-630-916A-74	Sequence 74, Appl
30	145	3.7	51	US-08-630-916A-73	Sequence 73, Appl
31	141	3.6	38	US-08-630-916A-36	Sequence 36, Appl
32	140	3.6	54	US-08-630-916A-118	Sequence 118, App
33	139	3.6	51	US-08-630-916A-116	Sequence 116, App
34	138	3.6	1105	US-08-710-249-2	Sequence 2, Appl
35	138	3.6	1105	US-08-220-157A-2	Sequence 2, Appl
36	133	3.4	224	US-08-630-916A-50	Sequence 50, Appl
37	132	3.4	51	US-08-630-916A-115	Sequence 115, App
38	130	3.3	38	US-08-630-916A-24	Sequence 24, Appl
39	130	3.3	38	US-08-630-916A-26	Sequence 26, Appl
40	130	3.3	38	US-08-348-518C-15	Sequence 15, Appl
41	130	3.3	38	US-08-348-518C-18	Sequence 18, Appl
42	130	3.3	38	US-08-476-509B-15	Sequence 15, Appl
43	130	3.3	38	US-08-476-509B-18	Sequence 18, Appl
44	130	3.3	335	US-08-844-312-2	Sequence 2, Appl
45	128	3.3	38	US-08-630-916A-25	Sequence 25, Appl

## ALIGNMENTS

RESULT 1  
US-08-539-205A-2  
Sequence 2, Application US/08539205A  
Patent No. 6001619

## GENERAL INFORMATION:

APPLICANT: Beach, David H.  
APPLICANT: Caligiuri, Maureen  
APPLICANT: Nefsky, Bradley  
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HONG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539, 205A  
FILING DATE: 04-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-539-205A-2

Query Match

Best Local Similarity 75.0%; Score 2913.5; DB 3; Length 735;

Matches 552; Conservative 63; Mismatches 71; Indels 71; Gaps 9;

Db 5 IKRLTVLCAKMLAKKDFRLPDPEAKIVVDGSGCHSTDTVKNLDPKMGHDLVYVK 64  
12 VKRLTVLCAKMLAKKDFRLPDPEAKIVVDGSGCHSTDTVKNLDPKMGHDLVYVK 58  
65 TDSITISWNNHKKIKKQAGFLGCVRLISNLSRLKDTGYRDLCKLNPEDTAVVGQ 124

```

Db 59 SDSVTISVNNHKKIKKQAGFLGCVRLLSNAINRLKDTGQRLDCLGENDNTVARGQ 118
QY 125 IVVSJQTRDRIGTGGVYVDCRGLBNE-----GTVY----- 155
Db 119 IVVSJQSRDRIGTGGVYVDCRGLBNE-----GTVY----- 178
QY 156 ---EDSGPRPLSCFMEEPAPYDSTGAAGGNCRFVSPQDQRLAQRANDVAGSL 213
Db 179 ASEYSSPGRPJSCFVDEMTPLSGINGATCG-----QSSDPLAARRRARSRRHNYM 229
QY 214 QTPQNRPHGQSPFLPEGEYQRTTVQGVYFLHTQTVSTWMDPRIPRDLNSVNCDELGP 273
Db 230 ---SRTHLHTPPDLPEGEYQRTTVQGVYFLHTQTVSTWMDPRIPRDLNSVNCDELGP 285
QY 274 LPPGWEVSTVSGRIYFVDHNNRTTQFTDPR---LHIMNHQCCLEKPSQPLPSPSGS 329
Db 286 LPPGWEIRNTATGRYFVDHNNRTTQFTDPRISANLHLYLRQQLKQOQOQVY---S 341
QY 330 L---EDELPAQRYERDLYOKLKYLRHLSLQOPAGHCRLEVSREEIFEESYQIMKMR 386
Db 342 LCPDTECLTVPRYRDLVQKLKIRQLSQQOPAGHCRLEVSREEIFEESYQIMKMR 401
QY 387 PKDLKRLMVKFRGEEGLDYGVAREMLYLLCHEMLNPPYGLFOYSTDNITMLOINPSS 446
Db 402 PKDLKRLMVKFRGEEGLDYGVAREMLYLLCHEMLNPPYGLFOYSTDDIYTLQINPDSA 461
QY 447 INPDLSYFHFYGRIMGAVFNGHYINGFTVPFYKOLGKPIQSLDESVDPELHKSIV 506
Db 462 VNPBHLSTFHFVGRIMGAVFNGHYIDGFTLPRFKOLGKSTLDDMELVDPDLHNSLV 521
QY 507 WLENDITPVLDHTECFVEHNAFGRILQHELKPNGANVPTEENKKEYRLLVYNNRPMRGI 566
Db 522 WLENDITGVLDHTECFVEHNAFGEIILQHELKPNKXSIPVNEENKKEYRLLVYNNRPMRGI 581
QY 567 EAOFLALQGFNEELIPOHLKPEPOKLELITIGLDKIDLMDKSNRIRKCVADSNIVR 626
Db 582 EAOFLALQGFNEELIPOHLKPEKLELITIGLDKIDVMDKSNRIRKCVADSNIVR 641
QY 627 WFOQAVETFEDEERRARLLQFTGSTRVPLQGFKALQSGTGAAGRLFTIHLIDANTNLP 686
Db 642 WFKWAVEFEDEERRARLLQFTGSTRVPLQGFKALQ---GAAGRLFTIHLIDANTNLP 696
QY 687 KAHTCFNRIDIPYBSYKLYEKLITAVEETCGFAVE 723
Db 699 KAHTCFNRIDIPYBSYKLYEKLITAVEETCGFAVE 735

RESULT 2
US-09-392-163A-2
; Sequence 2, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Calliguri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-2

Query Match 75.0%; Score 2913.5; DB 4; Length 735;
Best Local Similarity 72.9%; Pred. No. 7.1e-281;
Matches 552; Conservative 63; Mismatches 71; Indels 71; Gaps 9;

QY 5 IIRLTVLCANLAKKDFRLLDPFAKTYVDGSGGCHSTDTYKNTLDPKRNQHYDLYGK 64
Db 12 VKRLT-----GLPDPFAKTYVDGSGGCHSTDTYKNTLDPKRNQHYDLYGK 58
QY 65 TDSITISVNNHKKIKKQAGFLGCVRLLSNAISRLKDTGQRLDCLGENDNTVARGQ 124
Db 59 SDSVTISVNNHKKIKKQAGFLGCVRLLSNAINRLKDTGQRLDCLGENDNTVARGQ 118
QY 125 IVVSJQTRDRIGTGGVYVDCRGLBNE-----GTVY----- 155
Db 119 IVVSJQSRDRIGTGGVYVDCRGLBNE-----GTVY----- 178
QY 156 ---EDSGPRPLSCFMEEPAPYDSTGAAGGNCRFVSPQDQRLAQRANDVAGSL 213
Db 179 ASEYSSPGRPJSCFVDEMTPLSGINGATCG-----QSSDPLAARRRARSRRHNYM 229
QY 214 QTPQNRPHGQSPFLPEGEYQRTTVQGVYFLHTQTVSTWMDPRIPRDLNSVNCDELGP 273
Db 230 ---SRTHLHTPPDLPEGEYQRTTVQGVYFLHTQTVSTWMDPRIPRDLNSVNCDELGP 285
QY 274 LPPGWEVSTVSGRIYFVDHNNRTTQFTDPR---LHIMNHQCCLEKPSQPLPSPSGS 329
Db 286 LPPGWEIRNTATGRYFVDHNNRTTQFTDPRISANLHLYLRQQLKQOQOQVY---S 341
QY 330 L---EDELPAQRYERDLYOKLKYLRHLSLQOPAGHCRLEVSREEIFEESYQIMKMR 386
Db 342 LCPDTECLTVPRYRDLVQKLKIRQLSQQOPAGHCRLEVSREEIFEESYQIMKMR 401
QY 387 PKDLKRLMVKFRGEEGLDYGVAREMLYLLCHEMLNPPYGLFOYSTDNITMLOINPSS 446
Db 402 PKDLKRLMVKFRGEEGLDYGVAREMLYLLCHEMLNPPYGLFOYSTDDIYTLQINPDSA 461
QY 447 INPDLSYFHFYGRIMGAVFNGHYINGFTVPFYKOLGKPIQSLDESVDPELHKSIV 506
Db 462 VNPBHLSTFHFVGRIMGAVFNGHYIDGFTLPRFKOLGKSTLDDMELVDPDLHNSLV 521
QY 507 WLENDITPVLDHTECFVEHNAFGRILQHELKPNGANVPTEENKKEYRLLVYNNRPMRGI 566
Db 522 WLENDITGVLDHTECFVEHNAFGEIILQHELKPNKXSIPVNEENKKEYRLLVYNNRPMRGI 581
QY 567 EAOFLALQGFNEELIPOHLKPEPOKLELITIGLDKIDLMDKSNRIRKCVADSNIVR 626
Db 582 EAOFLALQGFNEELIPOHLKPEKLELITIGLDKIDVMDKSNRIRKCVADSNIVR 641
QY 627 WFOQAVETFEDEERRARLLQFTGSTRVPLQGFKALQSGTGAAGRLFTIHLIDANTNLP 686
Db 642 WFKWAVEFEDEERRARLLQFTGSTRVPLQGFKALQ---GAAGRLFTIHLIDANTNLP 696
QY 687 KAHTCFNRIDIPYBSYKLYEKLITAVEETCGFAVE 723
Db 699 KAHTCFNRIDIPYBSYKLYEKLITAVEETCGFAVE 735

```

RESULT 3  
US-08-539-205A-4  
Sequence 4, Application US/08539205A  
Patent No. 6001619  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Caligiuri, Maureen  
APPLICANT: Nelsky, Bradley  
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/539,205A  
FILING DATE: 04-OCT-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CSV-005.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 766 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-539-205A-4

Query Match 37.3%; Score 1471; DB 3; Length 766;  
Best Local Similarity 40.3%; Pred. No. 4.2e-137;  
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;  
QY 4 SIKRLTLVLCANLAKKDFRLPPFAKIVDGGGCHSDTVKNTLDPKMHYDLVYG 63  
DB 7 SRIKRTIIVADGLYKRDVFRFPDPAVLTVDS-QHTTTALIKTLNIPYMETFEVNT 65  
QY 64 KTDSTITSVMNKKIHKKQAGFLGCVRLISNAISRDKDTGYRL-DLCKLNPSSTDV 121  
DB 66 DNSTIALQVFDQKFF-KKKGQGLGVINLRYGDVLDLAIGDDMLTRDLKSN-ENTV 122  
QY 122 RQGIIVSLQ-----TPDRIGGGSVVOGRLLNEGTVYD 157  
DB 123 HGKIILNSTAOSTLOVPSSASGARTQRTSTINDPQSSKSSVSBNPSSASGPTD 182  
QY 158 SGPG-----PPLSCFMEB---PAPY---TDSITG-----AAA 182  
DB 183 NAPASPASSEPRFSSFEQYGLPFGWERRTDNLGRYYVDNHRSTTWIRPLISVA 242  
QY 183 GGNCRVSEPS-----QDQRLQGLRNPDYAGSLQTPNRRPHGHSPLPGYD 244  
DB 243 GAAALHSSASSANVTEGVQPSSSNARTRTESVLTSNATYAG-----SGELPGMB 296  
QY 235 RTTVQGVYFLHTQTVSTWHDPRIPRLNSVN-----CDLGLPDPGWEVRS 283  
DB 297 RYTBEGRPYFDNHTTRTTWDPDRQYIISYGGPNNAITQQCPVSGGLPGWEMRLT 356  
QY 284 VSGRIYVDHNNTQTQTPDLRHIMHQCCKLPSPPLPLPBGSLDEDELPAQYERD 343  
DB 357 NTAFVYFVDNHTKTITWDPR-----LPS--SL-DQNVF--QYKRD 392

QY 344 LVOKLTVLRHSLSQPQAGCRTEVSREELFEESYRQIMKMRPKDLKRLMYKFRGE 403  
DB 393 FRRKLIYFLSQPAL-HPLPGQCHIKVRNHHIFEDSVABIVRQSTDLKRLTKFDGE 451  
QY 404 LDYGVAREWLYLICHEMLNPYGLPQYSTDNIMYLQINPDSSINPDHLSYFHEVRIM 463  
DB 452 LDYGLSREYFLLSHMNPFCYCLFEYSVDNTYLQINSHSGINPEHLNFKRIGVIG 511  
QY 464 LAVFHGHYNGFTVPYRYKQLGKPIQLSDLESVDPLHSLVWLENDITPTVDHFCV 523  
DB 512 LAIFHRFVDAFPVASEYKILQKRVTLQWESMDASYSLWILNDITGVLDLFSV 571  
QY 524 EHNAFGRILQHELKPNRNVPVTEENKKEYRLVYNNRPMRGIEMACPLAQGFNELIPQ 583  
DB 572 EDNCFGEVYTLDKPNGRNIEVTEENKREYVDLTVW-IKRIIEQFNAPHEGSELIPQ 630  
QY 584 HLKPFQKLELLIIGLIDKIDLNDWKSNTRLKCVADSINVFQWAVETPDEERARL 643  
DB 631 ELINVFDERLELLIGISBIDMEDWKKKDYRSSENDQIKFWELMEWSNEKSRSL 690  
QY 644 LOFTGSTRVYLOGFKALQSGTGAAPRLFTHLIDANTDNLPAKHTCFNRIDIPY 703  
DB 691 LOFTTISRIVNGFKDQSD--GRKFTIEKA-GEPMKLPKATHCFRDLPLPTSK 746  
QY 704 EKYEKLTAVEETCGPAVE 723  
DB 747 KDLHKLSIAVEETIGFQGE 766

RESULT 4  
US-09-392-163A-4  
Sequence 4, Application US/09392163A  
Patent No. 6503742  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Caligiuri, Maureen  
APPLICANT: Nelsky, Bradley  
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/392,163A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/539,205  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CSV-005.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 766 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-392-163A-4  
Query Match 37.9%; Score 1471; DB 4; Length 766;

Best Local Similarity 40.9%; Pred. No. 4.2e-137;  
Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;

QY	51	STIRLPLVLCAXKLAKKDEFFRLJPDPAKTVWGSGGCHSTDPVKNXLDLPRKXNGHJDLVYG	63
Db	7	SRIRKRTIVAADGLYKRDVFRPPDPFAVLTVUGE-QHTTAIKKTLNPLYNNEFEVAVNT	65
QY	64	KTDSITISVWNAHKIKHKQAGFLGCVRLLSNAISRLKDTGYORL--DLCKLNSD7DAY	121
Db	66	DNSTIAIQVDPKKP-KKKGGGFLGYNLRVGDVLDLAIIGDEMTRDLKXSN--ENTVY	122
QY	122	RQGIIVSLQ-----TMDRIGTGSVYDCGGLLENGTYED	157
Db	123	HGKIILNSTAOSTLVQVPSAAGARTQRTSITVNDPOSSKSSVSRRNPASSRAGSPTRD	182
QY	158	SGCF-----RPLSCFME-----PAPY--TDSGT-----AAA	182
Db	183	NAPASAPASEPPTFSFEDQYGRLLPGMERITDUGTYVDHNTRESTTWRNLSVA	242
QY	183	GGGNCRFVSEPS-----QDRLQAOPLRNPVDYRGLQTPQNRPHGHSPELPEGYEQ	234
Db	243	GAAPAAELHSSASANTYEGVQESSMAARTASVLTNATAG-----SGELPPEMQ	296
QY	235	RTTVQGVQVYLNHTQGVSTWMBPRLPRDLSN-----CDELGPLPQMEVST	283
Db	297	RYTPEGEPYFVDNNTTTTWDPDRQOYTRSYGGENMATIQOQPVSGQPLPSGMEWRLT	356
QY	284	VSGRIYFVDNNRTTOFTDPRLHIMNHOCQKEPSOPLPLPSEGSLEDELPAPORYERD	343
Db	357	NTARVYFVDNNTTTTWDDPR-----LPS--SL-DQVVP--QYKRD	392
QY	344	LVQKLKYLRLHLSLQOPQAGHCRIRKSRBEIPEESRYQIMKRPDLKKRLMYKFRGEEG	403
Db	393	FRKRLIYFLSQPL--HPLPGQCHIKRNRNHIREDYAIMQSATDLCKRIMIFDEBDG	451
QY	404	LQVGVAREWLYLLEHMLNPYYGLFOYSTNDIYNLQINPSSINPDLSTYFHVGRIMG	463
Db	452	LQVGGLSREYFLHSHEMENRPPCLPEYSVUNYTLQINPHSGINPEHIAFKRIGVIG	511
QY	464	LAVFHGYINGGTFVFPYKQLGKPLQSLDSLESDVDELAHSLWILBENDITPVLDHTFCV	523
Db	512	LAFHFRFVDPAFVVSFYIMLQKVTLODMESMDAAYRSLVWILNDITGVLDLTFSV	571
QY	524	EHNAFGRILCHELKPGRNRPVTEENKKGYVALYNNRPFMGIEINQPLALOKGNELIPQ	583
Db	572	EDNCFEGVAVITIDKPGRNIEVTEENKRYVDLTVW--IQKRIEQQFAFHGSESLIPQ	630
QY	584	HLKPFPOKELIILIGLKDILDNMKSSTRYLKHCVAASNTVVRMFWQAVTFPDEERARL	643
Db	631	ELINVEDERLELILIGISEIDEMDKKHGDORSSENDQILIKFWELMDMSMEKSR	690
QY	644	LQVTVGSTRVPLGGFALOGSTGAAGPRLFTHLIDANTDULPRAHTCFNADIPYESTY	703
Db	691	LQFTTGSRIPVNGFDLOGSD---GPRKFTIEKA-GEPNKLPRAHTCFNLDLPYTSK	746
QY	704	EKLXKYLTAIVETGCPAVE	723
Db	747	KQDHDKLSIAVETIGFQGE	766

RESULT 5  
 US-09-070-060-4  
 Sequence 4, Application US/09070060  
 Patent No. 5976849  
 GENERAL INFORMATION:  
 APPLICANT: Hustad, Carolyn M.  
 APPLICANT: Childs, Carolynn M.  
 TITLE OF INVENTION: Human E3 Ubiquitin Proteinase  
 TITLE OF INVENTION: Ligase  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: ZENCA Pharmaceuticals, Inc.  
 STREET: 1800 Concord Pike

```

1      CITY: Wilmington
2      STATE: DE
3      COUNTRY: USA
4      ZIP: 19850-5437
5
6      COMPUTER READABLE FORM:
7      MEDIUM TYPE: Diskette
8      COMPUTER: IBM Compatible
9      OPERATING SYSTEM: DOS
10     SOFTWARE: FastSeq for Windows Vers
11
12     CURRENT APPLICATION DATA:
13     APPLICATION NUMBER: US/09/070,060
14     FILING DATE: 30-APR-1998
15
16     CLASSIFICATION:
17
18     PRIOR APPLICATION DATA:
19     APPLICATION NUMBER: 60/073,839
20     FILING DATE: 05-FEB-1998
21
22     ATTORNEY/AGENT INFORMATION:
23     NAME: Higgins, Patrick H
24     REGISTRATION NUMBER: 39,709
25     REFERENCE/DOCKET NUMBER: BHM.70312
26
27     TELECOMMUNICATION INFORMATION:
28     TELEPHONE: 302.886.4889
29     TELEFAX: 302.886.8221
30
31     TELEX:
32
33     INFORMATION FOR SEQ ID NO: 4:
34     SEQUENCE CHARACTERISTICS:
35     LENGTH: 854 amino acids
36     TYPE: amino acid
37     STRANDEDNESS: unknown
38     TOPOLOGY: unknown
39
40     MOLECULE TYPE: peptide
41
42     OS-09-070-060-4

```

Query Match	33.9%;	Score 1318;	DB 2;	Length 854;
Best Local Similarity	35.4%;	Pred. No. 8.9e-122;		
Matches 316;	Conservative 127;	Mismatches 229;	Indels 220;	Gaps 29

QY	6	KI RLATYCAK--NLAKDFRLLDDPPAKIVVDSQSCHSTDPKVTKATLDPKNHOAYDLYVGK	64
Dd	9	QLOITVYSKKLEKNKCMFQG--PSPVRYATVD--GQSKTEKCNNTNSPPKKQLOLYATVP	65
QY	65	TDSITITSVMHKKIHKXQAGFLGCVRL-----LSNAISLKDGTGYQR-----	107
Dd	66	TSKLCERVWSHQTLL--KSDVLLTAGAGDIYETLKSNNMKLEEVMTLOJVGKEPTEYM	122
QY	108	--LDLCUKLNSPTDAVRGQIVYS-----LQTRD--RIGTGGS--VYDCR	145
Dd	123	GDLSTVC---LDGLOVEAEVVTNGTSCSESTTQNDGCGTRDDTRVSTNGSEDPFYAAS	178
QY	146	G-----LLENEGTYVEDSGGRFLSCFMEEPAYT-----DSTG	179
Dd	179	GENKRANGNNSPBLNSCG--FKPSRP RP-----SRPPPTPRRPAVSNGSPSTNSDSG	231
QY	180	AA-----AGGNGRPFVESPQQ-----SQRLOAQR	203
Dd	232	SSGTSLPPNNINNVNTSTSEGAITSGLLPILTTISGSGSRPLATVSQALPBGMEQRV--	287
QY	204	LRNPVYRGSLL-----QTPONRPHGHQSBELPEGYEQRTTVGQGYFLPHQTGVSTWH	255
Dd	288	---DQHGKVVYVDHEVKRTTWDRPE-----FLPPGMEKRVDMNGRIYYVDHFTPTTWO	338
QY	256	DRIP-----RDL-----NSVNDDELGRPLPPGMEWR	281
Dd	339	RPLFESVRNVEQWOLQORSOLOGAMQOENQRFITYGNQULPATSOKEKFPDLGRLPPGMEKR	358
QY	282	STVSGRIYFVDHNNRTTQFTDPRLLHIMNHQCOJKEBSCPLP-----	324
Dd	399	TDSNGKVVYFVNHNRTLTQMEDPR-----SGGQLNE--KELPBGMEKRFIVDGIPIFYVDH	450
QY	325	-----PSSGSLDEDELPAQCYERDLYQKLYLA--HELSLOQPOAGHGRISRE	372
Dd	451	NRATATTYIDPRGKSGALDNGPQIAYVDRPKAKVYFRFQCOQLAMP-----HKTITVTRK	506

```

QY 373 EIFESYRQIMKMKPKDKLKRIMKFRGEGGLDYGVAREWLYLICHEMLNPPYGLFOYS 432
DB 507 TLFEDSFQOIMSPSPQDLRRRLWVIFPGESGLDYGVAREWFLLSHEVLANPWCLEFEYA 566
QY 433 TDNIYMLQINPDSSINPDHLSYFHFVGRIMGLAVFGHYINGFTVPFPYKQLGKPIQLS 492
DB 567 GKNVYCLQINPASYINPDHLYFRFGIFRIAMALFHGKFIIDTGFSLPFIKRLINKPVGLK 626
QY 493 DLESVDPRLHKSJLVILENDITPV-LDHTFCVENHAPGRILQHELKPNKRNAPVTEENK 551
DB 627 DLESIDPEFNSLSIWKENNIEECGLMEYFSVDKELIGIKSHDLKPNGNILVTEENKE 686
QY 552 EYRLVYVNRFRMGIEAOFALQKGFNELIPQHLKPFQKELIIGGLDKIDLNDWKS 611
DB 687 EYIRVAVMRSLRGVEEQTAFFEGFNEILPQYLYQYDALEVLGCMGEIDLNDWQ 746
QY 612 NTRLKHCVAADSNIVRFWQAVETPDEERRALLQFVGTSTVPLQGFALQSGTGAAGR 671
DB 747 HAIRYHYRTSKQIMFWQFVKEDINEKRMRLQFVGTGCLPVGGFADLMGNSN--GPQ 803
QY 672 LFTIHLIDANTDLPKATCFNRIDIPYSEYEKLYEKILTAVEETCGFAVE 723
DB 804 KFCIEKV-GKENWLPKSHTCFNRDLDPYKSYEQKELFLAIEETBEFGQE 854

```

## RESULT 6

```

US-09-357-746-4
; Sequence 4, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENNECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM.70312.N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PasteSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-357-746-4

```

```

Query Match 33.9%; Score 1318; DB 3; Length 854;
Best Local Similarity 35.4%; Pred. No. 8.9e-122;
Matches 316; Conservative 127; Mismatches 229; Indels 220; Gaps 29;

```

```

QY 6 KIRLVCAK-NLAKDPFRLDPFAKIVVDSGQCHSTDVKNLTDPKKNQHYDIYVGK 64
DB 9 QLGITVSAKLEKKNKMFQ-PSPYVEVTD--GQSKTEKKNNTNPKWKQPLVIVTP 65
QY 65 TDSITISVNAHKIKHKQAGFLCCVL-----LSNAISLKTGYQR----- 107
DB 66 TSKLCFRWVSHQTL--KSDVLTGTAGLDIYETLKSNNMKILEVVMTLQVGEKPEITM 122
QY 108 --LDLCKLNPEDTAVRGOIVS-----LQTRD--RIGTGS--VYDCR 145
DB 123 GDLVVC---LDGLOVEAEVVTNGETSCSESTQNDGCRTRDTRVSTNGSEPEVAAS 178
QY 146 G-----LIENBGTYEDSGPGRPLSCNFEAPPT-----DSTG 179
DB 179 GENKRVANGNSPSLSNGG--FKPSRPPRP-----SRPPPPPRRPAVNGSPSTNSDSDG 221
QY 180 AA-----AGGRCFVSPSCQ-----DQRLQQR 203
DB 232 SSTGSLPTNTNNTVSTEGATSGIILPLTISGSGSRPLNTVQAPLPQWMEKRV----- 287
QY 204 LRNPDPVRSGL-----QTPQNRPHGHQSPLEFGEYQRTTVGQVYFLHTQTGVSTWH 255

```

```

DB 288 ----DQGRVYVYVDHVEKRTWDRP-----PLPQWMEKRVYDNGKRIYYVDHFRITTTWQ 338
QY 286 DPRDP-----RL-----NSVNDDELGPRLPGMEVR 281
DB 339 RPTLESVANYBQWOLQBSQLOGAMQGNRFIYGNQDLFATSONKEFPDLPGLPQWMEKR 398
QY 282 STVGGRIFYVDHNNRTQFTDPRLHIMNHQCOLKEPSQPLPL----- 324
DB 399 TDSNCRVYFVVAHNRITQWEDPR-----SGQLNE--KPLPQWMEKRVTVDGIPYFVDH 450
QY 325 -----PEEGSLEDLELPAQRYERDLYQKLVLR--HELSLOQPOAGCHGIEVSR 372
DB 451 NRRATVYIDPRTGSAIDNGPOLAYVRDFKAKVQYFFRWCQOLAMPQ---HITVTRK 506
QY 373 EIFESYRQIMKMKPKDKLKRIMKFRGEGGLDYGVAREWLYLICHEMLNPPYGLFOYS 432
DB 507 TLFEDSFQOIMSPSPQDLRRRLWVIFPGESGLDYGVAREWFLLSHEVLANPWCLEFEYA 566
QY 433 TDNIYMLQINPDSSINPDHLSYFHFVGRIMGLAVFGHYINGFTVPFPYKQLGKPIQLS 492
DB 567 GKNVYCLQINPASYINPDHLYFRFGIFRIAMALFHGKFIIDTGFSLPFIKRLINKPVGLK 626
QY 493 DLESVDPRLHKSJLVILENDITPV-LDHTFCVENHAPGRILQHELKPNKRNAPVTEENK 551
DB 627 DLESIDPEFNSLSIWKENNIEECGLMEYFSVDKELIGIKSHDLKPNGNILVTEENKE 686
QY 552 EYRLVYVNRFRMGIEAOFALQKGFNELIPQHLKPFQKELIIGGLDKIDLNDWKS 611
DB 687 EYIRVAVMRSLRGVEEQTAFFEGFNEILPQYLYQYDALEVLGCMGEIDLNDWQ 746
QY 612 NTRLKHCVAADSNIVRFWQAVETPDEERRALLQFVGTSTVPLQGFALQSGTGAAGR 671
DB 747 HAIRYHYRTSKQIMFWQFVKEDINEKRMRLQFVGTGCLPVGGFADLMGNSN--GPQ 803
QY 672 LFTIHLIDANTDLPKATCFNRIDIPYSEYEKLYEKILTAVEETCGFAVE 723
DB 804 KFCIEKV-GKENWLPKSHTCFNRDLDPYKSYEQKELFLAIEETBEFGQE 854

```

## RESULT 7

```

US-09-070-060-3
; Sequence 3, Application US/09070060
; Patent No. 5976849

```

```

; GENERAL INFORMATION:
; APPLICANT: Husted, Carolyn M.
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENNECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: PasteSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H.
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHM.70312
; TELECOMMUNICATION INFORMATION:

```

/ TELEPHONE: 302.886.4889  
 / TELEFAX: 302.886.8221  
 / TELE: :  
 / INFORMATION FOR SEQ ID NO: 3 :  
 / SEQUENCE CHARACTERISTICS :  
 / LENGTH: 852 amino acids  
 / TYPE: amino acid  
 / STRANDEDNESS: unknown  
 / TOPOLOGY: unknown  
 / MOLECULE TYPE: peptide  
 / US-09-070-060-3

Query Match 33.5%; Score 1301; DB 2; Length 852;  
 Best Local Similarity 34.7%; Pred. No. 4,4e-120;  
 Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

```

QY 6 KRLTVLCAK-NLAKKDFRLPDPFAKIVVDGSGCHSTDTVKNLTDPKMNQHYDLYGK 64
DB 9 QLOIVTISAKLKENKKMFG-PSFYVEVTVD--GSKKTEKKNNTNSPKMKQPLVITVP 65
QY 65 TDSITISVNMKKIHKK--QGAGFLGCVRLSNALSRLKDT-----GYQR-----L 108
DB 66 VSKLHFRVWSHOTLKSVDLLGTAALDIYETLKSNNMKLEVVVTLQLOGDKPEPTTIGDL 125
QY 109 DLCKLNPEDTDAVRGOIVSLQT-----RDRICTGGS-----VYDC 144
DB 126 SIC-----LDGLQLEEVVNTNGETTSBSASQNDGSRKDETRVSTNGSDDEPDAGAGN 181
QY 145 RGL-----LENSTVYEDSGRPLSCFMEEPAPYDSTGAAGGNCRFVES----- 192
DB 182 RRVSGNNSPSLSNGG--FKPSRPPRP-----SRPPPPRRRPAASVNGSPSATSEDSGST 234
QY 193 -----PSODORLQAQRLNPD 208
DB 225 GSLPPTNTNTSSEATSGLIIPLTISGSGRPRLNPVTQAPLPPGMEQRY-----D 286
QY 209 VRGSL-----QTPQNRPHGHSPELPEGYEQRTTVQGVYFLHTQGVSTWHDPRIP 260
DB 287 QHGRVYVVDHVEKRTTWDRPE-----PLPPGMRERVDNMGRIVYVDHFTRTTWTQRPYLE 341
QY 261 -----RDL-----NSVNCDELGLPPGMEVRSYVSG 286
DB 342 SVRNTQWQLQSSQUGAMQCFNORFIYGNQDLFTSOSKSFDPGLPPGMEKRTDSNG 401
QY 287 RIYFVDHNNRTTOFTDPRLLHIMNHQCOLKEPSQPLP----- 324
DB 402 RYVFVNHNTRITQWMDPR-----SQGLNE--KRLPGMEMRFTVDSIPFYVDHNRRTT 453
QY 325 -----PSGSLDEDELPAPORYEDLVQKLKYLK--HELSLQOPQAGCRIVSSEELFEE 377
DB 454 TYIDRTGKSALDNGPQIAYVDRFKAQYFRFWCOQLAMPQ-----HIKIVTERTKLTLED 509
QY 378 SYROIIMKRPKDLKRLVWKFREGEGLDYGVAREMTYLLCHEMLNPFYGLFOYSTDNLY 437
DB 510 SFQQLMSSFPQDLRRRLVNIIFPGEGLDYGVAREMFLLSHEVLNMYCLFETAGKDNV 569
QY 438 MLOINPDSSINPDHLSYHFVGRIMGLAVFHGHYINGFTVPFYKQLGKPIQLSDLESV 497
DB 570 CLQINPASYINPDHLYKFRFGRIYALFEGKFIIDGFSLPFYKRIINLPVGLKDESI 629
QY 498 DPELHLSLVMLIENDITPV--LDHTEFCVENAFGRILQHELKPNRANVPVTEENKKEVRL 556
DB 630 DPEFNSILWKENNIEBCDLEMTFSDVKELIGIKHDKLPNGCNILVLENKEEYIRK 669
QY 557 YVNWEMKGIAGFLAQGFNELIPOLLKRPDQKLELIIIGLDKIDINDMKSNNRLK 616
DB 690 VAEWELSRGVEEQCAFEFEGNEILLPOQYLOYPDAKKELEVLICMQGSDINDMQRHAIR 749
QY 617 HCVADSNIVRWKQAVETFDSEERARLLQFTGSTRPLQCFKALQSTGAAGRLLTTH 676
DB 750 HYARTSKQIMFWQFVKEIDNEKMRLLQFTGTCRLPVGGFADLMKSN--GQKFCIE 806
QY 677 LIDANTDNLPRKATCFNRIDIPYSEYKELYEKLLTAVEETCGFAVE 723
  
```

DB 807 KY-GKENWLPBSHTCFNRIDLPYKSYEQLEKELFAIEETGFGQE 852

RESULT 8  
 US-09-357-746-3  
 / Sequence 3, Application US/09357746  
 / Patent No. 6087122  
 / GENERAL INFORMATION:  
 / APPLICANT: ZENEGA Limited  
 / TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE  
 / FILE REFERENCE: PHM 70312.N1  
 / CURRENT APPLICATION NUMBER: US/09/357,746  
 / CURRENT FILING DATE: 1999-07-21  
 / EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839  
 / EARLIER FILING DATE: 1998-02-05  
 / EARLIER APPLICATION NUMBER: US No. 608712209/070,060  
 / EARLIER FILING DATE: 1998-04-30  
 / NUMBER OF SEQ ID NOS: 15  
 / SOFTWARE: FastSeq for Windows Version 3.0  
 / SEQ ID NO 3  
 / LENGTH: 852  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 / US-09-357-746-3

Query Match 33.5%; Score 1301; DB 3; Length 852;  
 Best Local Similarity 34.7%; Pred. No. 4,4e-120;  
 Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;

```

QY 6 KRLTVLCAK-NLAKKDFRLPDPFAKIVVDGSGCHSTDTVKNLTDPKMNQHYDLYGK 64
DB 9 QLOIVTISAKLKENKKMFG-PSFYVEVTVD--GSKKTEKKNNTNSPKMKQPLVITVP 65
QY 65 TDSITISVNMKKIHKK--QGAGFLGCVRLSNALSRLKDT-----GYQR-----L 108
DB 66 VSKLHFRVWSHOTLKSVDLLGTAALDIYETLKSNNMKLEVVVTLQLOGDKPEPTTIGDL 125
QY 109 DLCKLNPSTDAVRGOIVSLQT-----RDRICTGGS-----VYDC 144
DB 126 SIC-----LDGLQLEEVVNTNGETTSBSASQNDGSRKDETRVSTNGSDDEPDAGAGN 181
QY 145 RGL-----LENSTVYEDSGRPLSCFMEEPAPYDSTGAAGGNCRFVES----- 192
DB 182 RRVSGNNSPSLSNGG--FKPSRPPRP-----SRPPPPRRRPAASVNGSPSATSEDSGST 234
QY 193 -----PSODORLQAQRLNPD 208
DB 225 GSLPPTNTNTSSEATSGLIIPLTISGSGRPRLNPVTQAPLPPGMEQRY-----D 286
QY 209 VRGSL-----QTPQNRPHGHSPELPEGYEQRTTVQGVYFLHTQGVSTWHDPRIP 260
DB 287 QHGRVYVVDHVEKRTTWDRPE-----PLPPGMRERVDNMGRIVYVDHFTRTTWTQRPYLE 341
QY 261 -----RDL-----NSVNCDELGLPPGMEVRSYVSG 286
DB 342 SVRNTQWQLQSSQUGAMQCFNORFIYGNQDLFTSOSKSFDPGLPPGMEKRTDSNG 401
QY 287 RIYFVDHNNRTTOFTDPRLLHIMNHQCOLKEPSQPLP----- 324
DB 402 RYVFVNHNTRITQWMDPR-----SQGLNE--KRLPGMEMRFTVDSIPFYVDHNRRTT 453
QY 325 -----PSGSLDEDELPAPORYEDLVQKLKYLK--HELSLQOPQAGCRIVSSEELFEE 377
DB 454 TYIDRTGKSALDNGPQIAYVDRFKAQYFRFWCOQLAMPQ-----HIKIVTERTKLTLED 509
QY 378 SYROIIMKRPKDLKRLVWKFREGEGLDYGVAREMTYLLCHEMLNPFYGLFOYSTDNLY 437
DB 510 SFQQLMSSFPQDLRRRLVNIIFPGEGLDYGVAREMFLLSHEVLNMYCLFETAGKDNV 569
QY 438 MLOINPDSSINPDHLSYHFVGRIMGLAVFHGHYINGFTVPFYKQLGKPIQLSDLESV 497
DB 570 CLQINPASYINPDHLYKFRFGRIYALFEGKFIIDGFSLPFYKRIINLPVGLKDESI 629
  
```

QY 498 DELHAKSLWILNDITPV-LDHTCEVHNAGRILOHELKNGNANVPTEENKEYVRL 556  
 Db 630 DEEFVNSLIMWKENNIEECDDLEMYFSDVDEILGELSHDLKPGNGAILTEENKEEYLRM 689  
 QY 557 YVNMREMGIEAOFALQKGFNELLIPQHLKPFDDKELELIGLDDKIDLNDMKNSTRK 616  
 Db 690 VAEWRLSRVEEOTQAFEGFNEILLPQYLQYFDAELEVLICGMQELDNDMQRHAIYR 749  
 QY 617 HCVADSNIVRWOAVETPDEERRARLLPVTGSTRVPLQGTALOGSTGAAGPRLFTIH 676  
 Db 750 HVAATSKQIMFWQFKEIDNEKRMRLLOFVTGTCRLPVGFGADLMGNSN---GPOKFCLE 806  
 QY 677 LLDANTDNLKPAATCFNRIDIPYSESEKYLEKTLAVEETCGFAVE 723  
 Db 807 KY-GKENMLPRSHTCFNRIDLPPYKSEYQKELFLAIBETGFGGE 852

## RESULT 9

US-08-895-601-6  
 ; Sequence 6, Application US/08895601  
 ; Patent No. 6060263  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beer-Romero, Peggy  
 ; APPLICANT: Strack, Peter J.  
 ; APPLICANT: Glass, Susan J.  
 ; APPLICANT: Rolfe, Mark  
 ; TITLE OF INVENTION: REGULATION OF KAPPA B (1KB) DEGRADATION,  
 ; TITLE OF INVENTION: AND METHODS AND REAGENTS RELATED THERETO  
 ; NUMBER OF SEQUENCES: 16  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109-2170  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/895,601  
 ; FILING DATE: 16-JUL-1997  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Vincent, Matthew P.  
 ; REGISTRATION NUMBER: 36,709  
 ; REFERENCE/DOCKET NUMBER: MITV-096.01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-832-1000  
 ; TELEFAX: 617-832-7000  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 927 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-895-601-6

Query Match 32.4%; Score 1259; DB 3; Length 927;

Best Local Similarity 33.6%; Pred. No. 7, 8e-116; Indels 240; Gaps 28;  
 Matches 307; Conservative 129; Mismatches 239;

QY 7 IRLTVLCARLAKKDFRLDPPEAKIV---VDGSGCHSTDTVAKTLDPKRN----- 55  
 Db 48 VRVAVIGAGIAGKDDIGASDPYRVVTLVDPMGVLTLSVQTKIKSLINKXNHEILLFRY 107  
 QY 56 ---OH-----YDLVYGKTD--SITISVWN-----HKIKKKOG 83  
 Db 108 HPOQHLLLEFVDENRLLTRDDQVDPVLPYLPTEPRRLRPTTFKDFVLRSHSKRV 167

QY 84 AGFLGCVALLSNALSRKDT-----GYQRDL-----LCKL----- 113  
 Db 168 KGVL-----FLKMTYLPKTSGEDDNDNAQAELEPGWVVLDDQPDACHLQOQCE 216  
 QY 114 -----NPSD--TDVARGQIVSLQ-----TRDR 134  
 Db 217 PSLPPGWEERQDILGRTYYVNHESRRTQWRPPTQDNLTDANGNLOQAPAFTRRQ 276  
 QY 135 IGTGGSVDRGULEN-----EGTYEDSG--PGRLSCFMEBPAYTD----- 176  
 Db 277 ISEETESVDNOSSESEMEITREDPAIYSSQAFSPSPSSNLDPYTHLABELNRLITFG 336  
 QY 177 ---STGAAGGNCR-----FVESP-----QDQRLQQL-- 204  
 Db 337 NSAVSOPASSNHSRRGSLQAYTFEERQPLFVLLPRTSSGLPGWEKQODERGRSYVDH 396  
 QY 205 -----RNPVRSLSIQPQ-----NRPHGQSPLEPEGYEQ 234  
 Db 397 NSRTTWTKPTVQAVETISQLTSSQSSAGPOSQASTDSGQOVTQPSLETQGFPLKGMV 456  
 QY 225 RTVQGVYFLATQGTGSTWEDPR--IPRLINSV---NDELGPLPPGMEVRSVSGRI 288  
 Db 457 RHAPNGRPFIDHTKTTTWEDPRLKIPALHKGTSLDTSDNLGFLPPGMEERTHTDGR 516  
 QY 289 YFVDHNNRTQFTDPRLLHIMHQCQLKEPQPLPLFSEGLDEBELPAQRYERDLVQL 348  
 Db 517 FYINNIKRTQWEDPRLNV-----ATGPAPV-----YSRDRYKX 553  
 QY 349 KVLRHESLQOPQAGHGRLEVSREIFESYRQIMKAPKD-LKKRLMVKFRGEGLDYG 407  
 Db 554 EFRKRLKQNDINPKFMKRLATVLEDSYRIMGYKRAFLARLMIERDEGKGLDG 613  
 QY 408 GVAREWLYLCHENLNPYYGLFQYSTDNLYWLQINPDSS-LNPHLSYFHVAGIMGLAV 466  
 Db 614 GVAREWFLISKEMENPYGLFEYSATDNYTLQINPNSGLCNEHLSYFKIGVAGAV 673  
 QY 467 FHGHYNGFTVPFQKLLGKPIQSLDLESVDPELHSLVWILNDITPVLDHTCFVCHN 526  
 Db 674 YHGKLDGFFTRPFRKMLHPITLDMESVDSSEYNSLRILNDPRL-IDLRFIDBE 732  
 QY 527 AFGRILOHELKNGNANVPTEENKEYVRLVNMREKNGIBAOFLAQKGFNELLIPQHL 586  
 Db 733 LFGQTHQHELKNGGSEIVVTKNKKKEYIVLVQWRFVNRIOKMAAFKGFELIPDLI 792  
 QY 587 KPFDQKELELIGLDDKIDLNDMKNSTRK-CVADSNIVRWOAVETPDEERRARLLQ 645  
 Db 793 KLFDENELELMCGAGDVADVMREHTYKNGYSANHQVIGWFKAVLMDSEKIRLLQ 852  
 QY 646 FVTGSTRVPLQGTALOGSTGAAGPRLFTIHLLDANTDNLKPAATCFNRIDIPYSESEK 705  
 Db 853 FVTGSTRVPMNGFALYGSN---GQOSTVWQW-GTPEKLRATHTCFRLLDPYSESE 908  
 QY 706 IYKLLTAVEETCGF 720  
 Db 909 LMDKLQMAIENTQGF 923

## RESULT 10

US-08-539-205A-6  
 ; Sequence 6, Application US/08539205A  
 ; Patent No. 6001619

; GENERAL INFORMATION:  
 ; APPLICANT: Beach, David H.  
 ; APPLICANT: Caligiuri, Maureen  
 ; APPLICANT: Nefsky, Bradley  
 ; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA

```

; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-539-205A-6

```

```

Query Match      31.6%; Score 1228.5; DB 3; Length 834;
Best Local Similarity 40.4%; Pred. No. 7.1e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

QY 93 LSNAISRL-----KDTGYQRLDCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
DB 186 LSEELSRRLQITPPDSNGQFSSLIQREPSRLRSCSVTDAVAEQGHLPPSVAYVHTTGG 245
QY 135 IGTGGSV-VDRCGLL-----ENEGTVYEDSGPRPLSCFMEEPAPYDSTGAAGGN-- 186
DB 246 LPSGWEERKDAKGRYYVNNHNRRTTWT-----RPIMQLAEDGA-----SGSATNSNNHL 295
QY 187 -----CRFVESPQDRLQARLRNPDVYGSU-----QTPQNRPHGQSP----- 227
DB 226 IEPQIRRRPRLSSPVTLSAPLEAGKQSPVRAVADTLSNPQSPQSPYNSPKQHKVTO 355
QY 228 --LPEGYQRTTVQGVYFHTQGVSTWHDRIIP-----RDLSVNCDELGPPLPGWEV 280
DB 356 SFLPGWEMRIAPNRPFIDHNKRTTWEDPRLKFPVHMSKSLNPNDLGPPPGWEE 415
QY 281 RSTVSGRIYFDHNHNRRTQFTDPRLRHIMNHQCQKXPSQPLPSEGSLEDELPQORY 340
DB 416 RIHLDGRFTFYIDHNSKITQWEDPRLN-----PALTG-----PAVPY 452
QY 341 ERDLVQKLVLRHLSLQOPQAGHCRLEVSREEIFEESYRQIMKM-RPKDLKRLMYKFR 399
DB 453 SRBFQKQYDFRKKLKKRPADIPNRFEMKLRHNNIFEESYRIRMSVKRPDLKARLWIEFE 512
QY 400 GEEGLDVGVARREMLYLICHEMLNPPYGLFOYSTDNLYMGINPDSGI-NPDHLSYHFV 458
DB 513 SEKGLDVGVARREMLFLSKEMENPYGLFEYSATDNTYLIQINNSGLCNEDHLSYFTFI 572
QY 459 GRINGLAVFHGYINGGTFVFPYKQLLQKPIQLSDLESDVPELHKSJLWILENDITPVD 518
DB 573 GRVAGLAVFHGKLDGFEIRFPFYKMLGKQITLNDMSVDSSEYNSJLWILENDITE 631
QY 519 HTFCVENAFGRILIOHLEKPKNGRNVPTSEENKKEKYVALYNNRPMRGLIEQFLALQKGN 578
DB 632 LMFCDIEBNFGQTYQVDLKRKNGSEIMVTNENKREYIDLVTQMRVNVNVOQKXNAFLGEGT 651
QY 579 ELIFQHLKPPDQKLELIIIGLKDILNDMKSNTRLKHCVADSN-IVRFMOAVERFDE 637
DB 692 ELIPIDILIKIDENELDELMLMGDLDVVDNMRQHSIKNQYCPRHPIQVFMXAVLMDA 751
QY 638 ERARLLOFTVGSIRVPLQGFKALQSGTGAAGPPLFIHLIDANTDNLPAHLTCNNRIDI 657
DB 752 EKRIIRLOFTVGTGRVPMNGFAELYGSN---GPQLFTIEQM-OSPEKLPRAHHCQFNRLD 807
QY 698 PPYESYEKLYEKLITAVEETCGF 720

```

```

DB 808 PPYTFEDLRKLLMAVENAQGF 830

```

```

RESULT 11
US-09-392-163A-6
; Sequence 6, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-6

Query Match      31.6%; Score 1228.5; DB 4; Length 834;
Best Local Similarity 40.4%; Pred. No. 7.1e-113;
Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

QY 93 LSNAISRL-----KDTGYQRLDCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
DB 186 LSEELSRRLQITPPDSNGQFSSLIQREPSRLRSCSVTDAVAEQGHLPPSVAYVHTTGG 245
QY 135 IGTGGSV-VDRCGLL-----ENEGTVYEDSGPRPLSCFMEEPAPYDSTGAAGGN-- 186
DB 246 LPSGWEERKDAKGRYYVNNHNRRTTWT-----RPIMQLAEDGA-----SGSATNSNNHL 295
QY 187 -----CRFVESPQDRLQARLRNPDVYGSU-----QTPQNRPHGQSP----- 227
DB 226 IEPQIRRRPRLSSPVTLSAPLEAGKQSPVRAVADTLSNPQSPQSPYNSPKQHKVTO 355
QY 228 --LPEGYQRTTVQGVYFHTQGVSTWHDRIIP-----RDLSVNCDELGPPLPGWEV 280
DB 356 SFLPGWEMRIAPNRPFIDHNKRTTWEDPRLKFPVHMSKSLNPNDLGPPPGWEE 415
QY 281 RSTVSGRIYFDHNHNRRTQFTDPRLRHIMNHQCQKXPSQPLPSEGSLEDELPQORY 340
DB 416 RIHLDGRFTFYIDHNSKITQWEDPRLN-----PALTG-----PAVPY 452
QY 341 ERDLVQKLVLRHLSLQOPQAGHCRLEVSREEIFEESYRQIMKM-RPKDLKRLMYKFR 399

```

Db 453 SREPKQYKPKKXKLPADINRFEMLKRNINFESEYRIMSVPDVLKARLMEFE 512  
 QY 400 GEEGLDYGVAEEMLYLCHENLNPYGLFOYSTDNITWLOINPDSST-NPHLSYFHFV 458  
 Db 513 SEKGIDYGVAAEWMFLLSKEMENPYGLFEYSADNDYTLQINPNSGCLNEDHLSYFTEI 572  
 QY 459 GIMGLAVHGHVINGFTVPFPKOLLGPKIQLSDLESVDPELHSLWILENDITPVLD 518  
 Db 573 GNVAGLAVHGHKLDLDFTRPFRYKMLKQITLNDMESVDSEYNSLKMILENDITE-LD 631  
 QY 519 HFFCEHNAFGRLLQELKPNGANVPVTEENKEVRLVYVNRPMKGLIQAFLAQKGFN 578  
 Db 632 LNFCDIENFGQYVYDCLKPNGSEIWMENKREYIDLVIQRFVNRVQKQNAFLLEGFT 691  
 QY 579 ELIPOLLPRPOKHELIIIGLKDIDLDMKSNTRKHCVADSN-IYVWFQAVETPE 637  
 Db 692 ELIPDLIKIPENELLELMCGLDVDVDMROHSIYKNGYCPNHVITQMFKAULMDA 751  
 QY 638 EERARLQFVTSSTVPLQGFKALQGSTGAAGPRLFTIHLIDANTDNEPKATCFNRIDI 697  
 Db 752 EKRIELQFVTGSRVPMNGFAELYSN---GPQLFTIEQW-GSPEKLPRAHTECNRLDL 807  
 QY 698 PPESEYKYEKLLTVEETCGF 720  
 Db 808 PPEYEFEDREKXLMVANAQGF 830

## RESULT 12

US-08-630-916A-48  
 ; Sequence 48, Application US/08630916A  
 ; Patent No. 601137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pirozzi, Gregorio  
 ; APPLICANT: Kay, Brian K.  
 ; APPLICANT: Fowlkes, Dana M.  
 ; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
 ; NUMBER OF SEQUENCES: 124  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: United States  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/630,916A  
 ; FILING DATE: 03-APR-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: MISROCK, S. LESLIE  
 ; REGISTRATION NUMBER: 18,872  
 ; REFERENCE/DOCKET NUMBER: 1101-203  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEFAX: (212) 896-8864/9741  
 ; INFORMATION FOR SEQ ID NO: 48:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 906 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: peptide  
 ; US-08-630-916A-48

Query Match 31.5%; Score 1225; DB 3; Length 906;  
 Best Local Similarity 35.5%; Pred. No. 1,86-112;  
 Matches 281; Conservative 115; Mismatches 227; Indels 169; Gaps 19;

QY 28 PFAKIVVDGSGQCHSDTIVKNTLDPKKNQHY-----DLYVGKTDISTITISVNNHKKIKKQ 83  
 Db 188 PNGSALTDGS-QLPBSRBSGTAVAPE-NRQOPSTNFGGRS-----RTRHSG 234  
 QY 84 AGFLGCVRLLSMALSRILKDTGYRLDICKLNPSTDVAHQIYVSLQTRD--IGTGGSV 141  
 Db 235 AS-----ARTPATGEQSPGASRRHQPKNSGHS 264  
 QY 142 VDCRGLLENSTGYED-----SGPRPLSCFEE-----PAYDSTGAAGGN 186  
 Db 265 GLANGVNDSEFTATDTEBBSVGVTSPPAPLBSVTNPNTSIPAPATAEG----- 317  
 QY 187 CRFVESPQD--QRLQARLRMPDVRGSLQTPONRPHGOSPELPE----- 230  
 Db 318 ---EEBSTGTOQLPA-----AAQAPDALPAGMBQRELPNGRVVYVDHTKTTT 363  
 QY 231 ---GYEORTVQGVYFLHTQGVSTVMDPRIRPDIN----- 264  
 Db 364 WERPLPQWEMERTDPRGRFVYVDHTRTTTWQRTAERYVRYEQMSQKQLOGAMGHS 423  
 QY 265 ---SVNDELGLPLPPGWEVRSVSGRYFVDHNNRTTQFTDPR----- 304  
 Db 424 QRLYQFWSASTDHDPLGLPPGWEKQD--NGRVYVNNHTRTTQWEDPRQGMWICPAL 482  
 QY 305 -----LHINNHQCOQLKEPSQPLPLBSGSLDEBELPAQYERDLYQKLVLR 352  
 Db 483 PPGWEMKYTESEGVYFVDHNTRTTTFPDPRPGFESGTQKQSPGAYDRSFRKXHQFFELC 542  
 QY 353 HELSIQFQACHCEIYSREIPEESYKQIMKRPXDKRLMYKFGEEGLDYGVAE 412  
 Db 543 HSNAL---PSHVKISVRQTLFEDSFQIIMMKPDLRRLIYIMGEBGLDYGIAE 598  
 QY 413 WLYLCHENLNPYGLFOYSTDNITWLOINPDSINPDHLSYFHFVGRINGLAVHGHYI 472  
 Db 599 WFFLSHEVLNPMVCLFEYAGKNVYCLQINPASSINPDHLYFFIGRIFIMALYHGHFI 658  
 QY 473 NGGFVFPYKOLLGPKIQLSDLESVDPELHSLWILENDITPV-LDHTFCVENHAFRI 531  
 Db 659 DTGFTLPPYKMLNRPPLTKDLESIDEPFNSIWIKENNIEBGCLELYFQDWEILQKV 718  
 QY 532 LQHEIKPNRNVVYTEENKEVRLVYVNRPMKGLIQAFLAQKGFNELLBQHLKPPDQ 591  
 Db 719 TTHELKEGGEISRYTEENKEVIMLTDWRTRGVEBQTKAFLQGFVAVLEWLRYPDE 778  
 QY 592 XELELITGLDKLIDNPKSTRKLKHCYADSNIYVWFQAVETDEERARLLQFVGST 651  
 Db 779 KELEIMLCQMOEIDMSDWQKSTIYRHYTKSKQIQWFQVQVKEMDNEKRIRLQFVGTG 838  
 QY 652 RYVLOGFKALQGSTGAAGPRLFTIHLIDANTDNEPKATCFNRIDIPPESEYKYEKLL 711  
 Db 839 RLPVGGFPAELYSN---GPQKFCIDKVKET-WLPRSHTCNRLDLPYKSYEQRLREKL 894  
 QY 712 TAVEETCGFAVE 723  
 Db 895 VAIEETEGGQE 906

## RESULT 13

US-08-630-916A-46  
 ; Sequence 46, Application US/08630916A  
 ; Patent No. 601137  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Pirozzi, Gregorio  
 ; APPLICANT: Kay, Brian K.  
 ; APPLICANT: Fowlkes, Dana M.  
 ; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
 ; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
 ; NUMBER OF SEQUENCES: 124  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Pennie & Edmonds  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York

STATE: New York  
COUNTRY: United States  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,916A  
FILING DATE: 03-APR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MISROCK, S. LESLIE  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-203  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 896-8664/9741  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 683 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-630-916A-46

Query Match 27.0%; Score 1047; DB 3; Length 683;  
Best Local Similarity 43.8%; Pred. No. 6,1e-95;  
Matches 217; Conservative 81; Mismatches 146; Indels 52; Gaps 10;

QY 190 VESPSQDRLQAKRNDVAGSLQ-----TPQNRHGQSPBLPGYHQT 236  
DB 225 MESVAFPEQWOSCR---NQLQAMQOFNQRVLVSASMLAENDVYG---PLPGMEKRV 277  
QY 237 TVCGQVYELTQTGVSTWHDPRIDPLNSVNCDELGLPGWEVRSVSGRIYFVDHNR 296  
DB 278 DSTBRYVFNNTKTQWEDPRGLONE-----EPLPGMEIRYRREGVRFVDHNR 331  
QY 297 TTQFTDPLHHIHHQCKEKSQPLPLPSEGLSEDELPQORYERDVLQKLVRLHLS 356  
DB 332 TTFKDPDPR---NGKSSV-----TKG-----POLAVERGFRWKLAFRY-LC 369  
QY 357 LQOQAGCRLEVSREEIFEESRYQIMKMRPKDKKRLMKFRGEEGLDYGVARRELYL 416  
DB 370 QSNALPSVAKINVERQTFEDSFQIMALKFYDLRRLYIFRGSGLDYGGLAREKFFL 429  
QY 417 LCHENLNPYGLFOYSTDNIMYLQINPDSSINPDHLSYFHEVGRIMGLAVFHGHYINGF 476  
DB 430 LSHVNLNPMYCLFEYAGKNNYCLQINPASTINPDHLSYFCFGRFIMALLFHGKFDITGF 489  
QY 477 TVPEYKQLGKPIQLSDLESVDPELHKSIVLILENDITPV-LDHTFCVENNAFGRIQHE 535  
DB 490 SLPEYKRLSKKLTIKDESIDTFEYNSLWIRNNNIECGLEVFSDMEILGKVTSHD 549  
QY 536 LKPNRNPVTEENKKEEYVLLYVMRMFGIEAFLQKGFENELLPOHLLKFPDQKELE 595  
DB 550 LKLGGSNLTVEENKDEYIGLMTWRSPRGQDETKAFLDGPNFVAVLQWLQFDEKLE 609  
QY 596 LIIGGLDKLIDNKSATRLKHCYADSNIVAMFQAVETPDERRALLOFTVGSRTVPL 655  
DB 610 VMLGQMEVDLADQRYVYRHVYTRNSKQIIFWQFVKTENDNEVRMLLOFTVGTGRPL 669  
QY 656 QGFALOGSTGAAGR 671  
DB 670 GGFALMGSN---GPR 682

RESULT 14  
US-08-247-904B-8  
; Sequence 8, Application US/08247904B  
; Patent No. 5981699

GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Eckstein, Jens W.  
TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley, Hoag & Eliot  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/247,904B  
FILING DATE: 23-MAY-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 874 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-247-904B-8

Query Match 13.7%; Score 530.5; DB 2; Length 874;  
Best Local Similarity 26.1%; Pred. No. 2.7e-43;  
Matches 190; Conservative 117; Mismatches 262; Indels 159; Gaps 26;

QY 98 SRLKDTGYRDLCKLNPSTDT---DAVRGOIVSLQTRDRIGTG-----GSVDCR 145  
DB 202 SRIGDSQGDNNLQGLGPDVSDVIDAIR-RVYTRILSNEXIEFALFVALYLSPNVGC 260  
QY 146 GLENEGVY-----EBSGPERPLSCMEBAPYTTDSTGAAG 183  
DB 261 LTYHN---VYSRDPYVNLFTIGMENRNHSPYELMALPLFCRAMSKLP-----LAAQ 311  
QY 184 GGNCRFVSPSQDQ-----RLOQRLRPD----- 208  
DB 312 GKLILNMRKNVADQRRMMEYFQQLITYKVISNEFNSNLVNMENSRNLVNDDAIVAAS 371  
QY 209 -----VRGSLQTPQRRPHGOS-PELP-----GYQRTTVQGVYFLHTQT 249  
DB 372 KCLRMVYVAVVGGVDVTDNHNEDDEEPIPSSEETLQELLEGERRRNKGRLVAPLETEL 431  
QY 250 GVST--WMDPRIPR--LNSVNCDELGLPGWEVRSVSGRIYFVDHNRRTQFTPRL 305  
DB 432 GYKTIIDCKRPLIPREFINE-----PLNEVLENDKYT--FKVRENKFSMTCPFI 482  
QY 306 HHNNHQCKEKSQPLPLPSEGLSEDELPQORYERDVLQKLVRLHLSLQOPQAGHC 365  
DB 483 INAVTKNGLYDNR-IMYSE-----RRTVL-YSLVQGOQLNPYL 522  
QY 366 RIEVSRREIFEES---YQIMKMRPKDKKRLMKFRGEEGLDYGVARRELYLICHEML 422  
DB 523 RLKVRBDHIIDALVRLLEMENPADIKQLYVEFEGSGVDGSGKFEFQAVEEIF 582  
QY 423 NPVYGLFOYSTDNIMYLQINPDSSINPDHLSYFHEVGRIMGLAVFHGHYINGGFTVPEYK 482  
DB 583 NPDIGMFTYD-ESRKLFWFNPSFP---ETESQFTLLIGIVGLALYNNCILDVHPMNVYR 638

QY 483 OLLGKPIQLSDLESVDPELHKSIVWILE--NDITPVLDHTFCV-EHNAFGRILOHELKEN 539  
 DB 639 KLMGKGLFVLDGSDHPIVLYOSIKDLLEYGVNVEDDMITTFQISQTNLFQNMATDLEN 698  
 QY 540 GRNVPTVEENKEEYRLVYVMRFRMGIEAQLALOKGF---NELIPQHLKPFDOKELE 595  
 DB 699 GKKIPITNENKEFVNLISDYILINKSVEKQFARRGFMVNTNESPILKYLFRP---EEIE 755  
 QY 596 LIIGGLDKIDLNDKMSNTRLK-HCVADSNIVRMQAVETPDEERRARLLQFTGSTTRVP 654  
 DB 756 LLIIGSRNLDFOALBETTEYDGYTRDSVLIREFWEIVHSFTDEQKRLFLQFTTGIDRAP 815  
 QY 655 LOGFKALOGSTGAAGPRLFTIHLIDANTDNLPKATCNFRIDIPYSEYKLYEKLITAV 714  
 DB 816 VGLGLKTKMIIAKNGP-----DTERLPTSHTCFNVLLPEYSEKELKERLKA 865  
 QY 715 EETCGFAV 722  
 DB 866 TYAKGFGM 873

## RESULT 15

US-08-767-942A-21  
 ; Sequence 21, Application US/08767942A  
 ; Patent No. 6068982

## GENERAL INFORMATION:

APPLICANT: Rolfe, Mark  
 APPLICANT: Chiu, M. Isabel  
 APPLICANT: Berlin, Vivian  
 APPLICANT: Dargatz, Veronique  
 APPLICANT: Draetta, Giulio  
 APPLICANT: Guillaume, Cottarel  
 TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: FOLEY, HOAG & ELLIOT LLP  
 STREET: One Post Office Square  
 CITY: Boston

STATE: MA  
 COUNTRY: USA  
 ZIP: 02109-2170

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/767,942A  
 FILING DATE: 17-DEC-1996

## ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.  
 REGISTRATION NUMBER: 36,709  
 REFERENCE/DOCKET NUMBER: MIV-029.04  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-832-7000  
 TELEFAX: 617-832-1000

## INFORMATION FOR SRO ID NO: 21:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 874 amino acids  
 TYPE: amino acid  
 TOPOLOGY: 1linear  
 MOLECULE TYPE: protein  
 US-08-767-942A-21

Query Match 13.7%; Score 530.5; DB 3; Length 874;  
 Best Local Similarity 26.1%; Fred. No. 2.7e-43;  
 Matches 190; Conservative 117; Mismatches 262; Indels 159; Gaps 26;

QY 98 SRLKDTGYRLDLCKLNPST---DAVRGQIVSLQTRDRIGTG-----GSVVDGR 145  
 DB 202 SRIDSSGSGDNKIQGLGSDVDVSDIDAIR-RVYTRILSNKEIETAFILNALVYLSPNVECD 260  
 QY 146 GLIENEGTV-----EDSGRGRLSCMEBPAPYTDSTGAAG 183

DB 261 LTYHN---VYSRDPNYINLFIIGMENENLHSPYLEMALDLFCAMSKLP-----LAAQ 311  
 QY 184 GGNCRFVESPQDQ-----RLOQRRLNPD----- 208  
 DB 312 GKILRLMSKINADIRMMETFOQLITYKVISNEFNSRLVNEANSNLVNDDAIYAA 371  
 QY 209 -----VRGSLQTPQNRPHQOS-PELPE-----GYEQTVOQGVYFLHTQT 249  
 DB 372 KCLMWVYANAVGGEVDTNHNEDEDEPIPESSSELTIOELLGERRNKKGLRVLPLETEL 431  
 QY 250 GVST--WEDFRIPRD--LNSNCELEPRLPGWVRSTVSGRIYFVHNHRTTOFTPRL 305  
 DB 432 GVKTLDCRKLPIPEEFINB-----PLNEVLMDKDY--PFKVTENKFSMTCPFI 482  
 QY 306 HHIMNHQCOLKEPQPLPLPESGSLDEDELPAQRYERDLVQKLVRLHLSLOQPOAGHC 365  
 DB 483 LNAVTKNLGLYYDNR-IRMYSE-----RRITVL-YSLVQGGQLNPYL 522  
 QY 366 RIEVSREEIFEES---YKQIYKMRPKDKRLMYKFRGEBGLDYGVARENLYLICHEML 422  
 DB 523 RLKYVRDHIIDDAVRLMEMIAMENPADLKKQLYVEFBEGQVDGVSKEFFQLVEBIF 582  
 QY 423 NPYVGLPQYSTDNIVYMLQINPDSSINPDHLSYFHFVGSIMGLAVFHGYINGCFVPEYK 482  
 DB 583 NPDIGMFTYD-ESTYLFMFNPSSF---ETEGQFTLIGIVGLAIYNNCILDVHFPMVYR 638  
 QY 483 OLLGKPIQLSDLESVDPELHKSIVWILE--NDITPVLDHTFCV-EHNAFGRILOHELKEN 539  
 DB 639 KLMGKGLFVLDGSDHPIVLYOSIKDLLEYGVNVEDDMITTFQISQTNLFQNMATDLEN 698  
 QY 540 GRNVPTVEENKEEYRLVYVMRFRMGIEAQLALOKGF---NELIPQHLKPFDOKELE 595  
 DB 699 GKKIPITNENKEFVNLISDYILINKSVEKQFARRGFMVNTNESPILKYLFRP---EEIE 755  
 QY 596 LIIGGLDKIDLNDKMSNTRLK-HCVADSNIVRMQAVETPDEERRARLLQFTGSTTRVP 654  
 DB 756 LLIIGSRNLDFOALBETTEYDGYTRDSVLIREFWEIVHSFTDEQKRLFLQFTTGIDRAP 815  
 QY 655 LOGFKALOGSTGAAGPRLFTIHLIDANTDNLPKATCNFRIDIPYSEYKLYEKLITAV 714  
 DB 816 VGLGLKTKMIIAKNGP-----DTERLPTSHTCFNVLLPEYSEKELKERLKA 865  
 QY 715 EETCGFAV 722  
 DB 866 TYAKGFGM 873

Search completed: September 21, 2004, 07:49:17  
 Job time : 28.106 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:28:49 ; Search time 92.5466 Seconds  
(without alignments)  
2285.667 Million cell updates/sec

Title: US-10-009-945-4

Perfect score: 4038

Sequence: 1 MSNPGRRNGPKRLTVLC.....EKLYEKLTALPETGFAVE 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*\n2: geneseqp1990s:\*\n3: geneseqp2000s:\*\n4: geneseqp2001s:\*\n5: geneseqp2002s:\*\n6: geneseqp2003as:\*\n7: geneseqp2003bs:\*\n8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4038	100.0	748	AAB31477	Aab31477 Amino aci
2	4031	99.8	748	AAW13384	Aaw13384 Human pro
3	4027	99.7	804	AAU19610	Aau19610 Human dla
4	3001	74.3	722	AAE33724	Aae33724 E3 ubiqui
5	3001	74.3	723	AAB31476	Aab31476 Amino aci
6	2992	74.1	722	AAW79861	Aaw79861 Human pro
7	2775	68.7	514	AAU87301	Aau87301 Novel cen
8	2173	53.8	1035	ABB61120	Abb61120 Drosophi
9	1809	44.8	335	ADC64238	Adc64238 Human SMT
10	1545	38.3	766	AAW13385	Aaw13385 Human pro
11	1530.5	37.9	832	ABP73459	Abp73459 Candida a
12	1453.5	36.0	869	ABJ26104	Abj26104 Aspergill
13	1387	34.3	975	AAW93167	Aaw93167 Human ZGG
14	1370.5	33.9	911	ADB75455	Adb75455 Prostata
15	1365.5	33.6	724	AAW35797	Aaw35797 Novel hum
16	1355.5	33.6	725	ADB49320	Adb49320 Novel hum
17	1354.5	33.5	854	ADB75453	Adb75453 Prostata
18	1354.5	33.5	927	AAV25170	Aav25170 Human KIA
19	1354.5	33.5	927	AAE33725	Aae33725 Human NED
20	1353.5	33.5	995	AAE33719	Aae33719 KIAA0439
21	1353.5	33.5	995	ADB75451	Adb75451 Prostata
22	1353.5	33.5	995	ADAE6001	Adae6001 Human pro
23	1353.5	33.5	995	ADAE4562	Adae4562 Human pro
24	1353	33.5	923	ABG01080	Abg01080 Novel hum
25	1348	33.4	834	AAU77715	Aau77715 Diseased

26	1340.5	33.2	759	6	ABJ25504	Abj25504 Aspergill
27	1336.5	33.1	854	6	ABP58332	Abp58332 Human cel
28	1336	33.1	854	2	AAV30949	Aav30949 Murine E3
29	1333	33.0	949	4	ABBS9631	Abbs9631 Drosophi
30	1333	33.0	949	4	ABBS9856	Abbs9856 Drosophi
31	1329	32.9	852	2	AAV30948	Aav30948 Human B3
32	1313.5	32.5	903	6	ABR64210	AbR64210 Angiogene
33	1301	32.2	739	4	ABG16477	Abg16477 Novel hum
34	1301	32.2	739	4	AAE68173	Aae68173 Atrophin-
35	1301	32.2	739	6	AAE32720	Aae32720 Atrophin-
36	1301	32.2	739	6	ABR41097	AbR41097 Human atr
37	1301	32.2	739	7	ADB96717	Adb96717 Human atr
38	1301	32.2	739	7	ADE82513	Ade82513 Human pro
39	1301	32.2	752	8	ADE77070	Ade77070 Human pro
40	1298.5	32.2	870	4	AAE05495	Aae05495 Human ubi
41	1298.5	32.2	870	6	AAE33722	Aae33722 Nedd-4-1i
42	1298.5	32.2	906	7	ADB49242	Adb49242 Novel hum
43	1297	32.1	922	4	ABBS0049	Abbs0049 Human hom
44	1297	32.1	922	5	ABBS0708	Abbs0708 Human sig
45	1285.5	31.8	906	2	AAW36795	Aaw36795 Novel hum

## ALIGNMENTS

RESULT 1  
AAB31477  
ID AAB31477 standard; protein; 748 AA.  
XX  
AC AAB31477;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Amino acid sequence of a human Smurf2 polypeptide.  
XX  
KW Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP;  
KW transforming growth factor-beta; human; TGF-beta; chondrogenesis;  
KW osteogenesis; blood differentiation; cartilage formation; hair growth;  
KW neural tube patterning; retinal development; heart induction;  
KW morphogenesis; tooth formation; gamete formation.  
XX  
OS Homo sapiens.  
XX  
PN WO200077168-A2.  
XX  
PD 21-DEC-2000.  
XX  
PF 12-JUN-2000; 2000WO-US016250.  
XX  
PX 11-JUN-1999; 99US-0138969P.  
XX  
PY (UNIV ) UNIV NEW YORK STATE RES FOUND.  
XX  
PA (HSCR-) HSC RES & DEV LP.  
XX  
PI Thomsen GH, Wrana J;  
XX  
DR MPI: 2001-071267/08.  
XX  
NR N-PSDB; AAF24853.  
XX  
PT Novel isolated Smurf protein useful for inhibiting bone morphogenic  
PT protein or tumor growth factor-beta activation pathway, for treating  
PT cancer and to block osteogenesis, hair growth, tooth formation.  
XX  
PS Claim 10; Fig 12; 107pp; English.  
XX  
CC The present sequence represents a human Smurf2 polypeptide. The  
CC specificatation also describes a Smurf1 polypeptide. Smurf polypeptides are  
CC negative regulators of Smad signal transduction, and antagonists of bone  
CC morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta)  
CC signaling pathway. Expression of Smurf1 in a cell is useful for  
CC inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf  
CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood  
CC differentiation, cartilage formation, neural tube patterning, retinal

CC development, heart induction and morphogenesis, hair growth, tooth  
 CC formation, gamete formation and a wide variety of tissue and organ  
 CC formation processes, and hinder the regeneration, growth, maintenance,  
 CC etc., of bone and other tissues that are dependent on the BMP pathway.  
 CC The polypeptide is useful for screening for various drugs and/or  
 CC antibodies that can either enhance the BMP pathway, or inhibit it  
 XX

XX Sequence 748 AA;

Query Match 100.0%; Score 4038; DB 4; Length 748;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 748; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSNPGRRNGPVKRLITVLCANLVKDFPLPPFAKVVVDGSGQCHSTDTVKNLDPK 60
DB 1 MSNPGRRNGPVKRLITVLCANLVKDFPLPPFAKVVVDGSGQCHSTDTVKNLDPK 60
QY 61 WNOHVDLYIGKSDSYTISVWNNHKKIKKKOGAGFLGCVRLSNAINRLKDTGYORLDCKL 120
DB 61 WNOHVDLYIGKSDSYTISVWNNHKKIKKKOGAGFLGCVRLSNAINRLKDTGYORLDCKL 120
QY 121 GPNNDIVRGQIVVSLQSRDRIIGTGGVYVDCSRLFDNDLPDGEERRTASGRIOYLNHIT 180
DB 121 GPNNDIVRGQIVVSLQSRDRIIGTGGVYVDCSRLFDNDLPDGEERRTASGRIOYLNHIT 180
QY 181 RTQWERPTRPASESSSPGRPLSCFVDENPTISGTNATCGSSDPLARRVRSQHRN 240
DB 181 RTQWERPTRPASESSSPGRPLSCFVDENPTISGTNATCGSSDPLARRVRSQHRN 240
QY 241 YMSRTHLHTPPDLDEGEYQRTTQGGVYFLHTQTGVSFWHDPRPRDLNSINCEELGFLP 300
DB 241 YMSRTHLHTPPDLDEGEYQRTTQGGVYFLHTQTGVSFWHDPRPRDLNSINCEELGFLP 300
QY 301 PGWEIRNTATGRVYFVDDNNRTTQTPRISANHLVLNQNOLKDOQQQVYVSLCPDDT 360
DB 301 PGWEIRNTATGRVYFVDDNNRTTQTPRISANHLVLNQNOLKDOQQQVYVSLCPDDT 360
QY 361 ECLTVPRYKRDLYOKLILRQELSQOQPOAGHCRIEVSRREIPEESYROYMKRXPDLWK 420
DB 361 ECLTVPRYKRDLYOKLILRQELSQOQPOAGHCRIEVSRREIPEESYROYMKRXPDLWK 420
QY 421 RLMKFRGEBGLDYGVAREWLYLSHEMLNPYGLFQYSRDDIYTLQINPDSAVNBEHL 480
DB 421 RLMKFRGEBGLDYGVAREWLYLSHEMLNPYGLFQYSRDDIYTLQINPDSAVNBEHL 480
QY 481 SYFFHVGIRMGMAVFHGHYIDGFTLPFYKQLGKSTITLDDMELVDDPLHNSLVWILEND 540
DB 481 SYFFHVGIRMGMAVFHGHYIDGFTLPFYKQLGKSTITLDDMELVDDPLHNSLVWILEND 540
QY 541 ITGVLDTFCVENHAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVWRFLEAGIEAOFIA 600
DB 541 ITGVLDTFCVENHAYGEIIQHELKPNKSIIPVNEENKKEYVRLYVWRFLEAGIEAOFIA 600
QY 601 LQKGFNEVIRPQHLKTDKEKELEIIIGLGIKIDNMKQNTRLKCTPDSNIYKMWKAV 660
DB 601 LQKGFNEVIRPQHLKTDKEKELEIIIGLGIKIDNMKQNTRLKCTPDSNIYKMWKAV 660
QY 661 EFPDEERRARLLQFVTSSSEVPLOGFYALQGAAPRLFTTHQIDACTNNLPKATCFENRI 720
DB 661 EFPDEERRARLLQFVTSSSEVPLOGFYALQGAAPRLFTTHQIDACTNNLPKATCFENRI 720
QY 721 DIPYSEYKLYEKLILAIETCGFAVE 748
DB 721 DIPYSEYKLYEKLILAIETCGFAVE 748

```

RESULT 2

AAW13384 ID AAW13384 standard; protein: 748 AA.

AC AAW13384;

XX 10-JUL-1997 (first entry)

XX Human protein ubiquitin ligase pub1.  
 DE  
 XX  
 XX Protein ubiquitin ligase; pub1; cdc25 phosphatase; CDK kinase; p53;  
 XX cell cycle; transgenic animal.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX MO9712962-AL  
 XX  
 XX 10-APR-1997.  
 PD  
 XX  
 XX 04-OCT-1996; 96WO-US015930.  
 XX  
 XX 04-OCT-1995; 95US-00539205.  
 XX  
 XX (COLD-) COLD SPRING HARBOR LAB.  
 XX  
 XX Beach D, Caligiuri M, Nefsky B;  
 XX  
 XX WPI; 1997-226206/20.  
 DR  
 XX  
 XX N-PSDB; AAT47040.  
 DR  
 XX  
 XX Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved in  
 PT cdc25 phosphatase and p53 ubiquitination, and regulate cell growth and  
 PT proliferation.  
 PT  
 XX  
 XX Claim 1: Page 74-77; 108bp; English.  
 PS  
 XX  
 XX

Human protein ubiquitin ligase pub1 (AAW13384), pub2 (AAW13385) and pub3 (AAW13386) are homologues of fission yeast pub1 (AAW13387) and were identified from cDNA clones (AAT47040-42) obtd. e.g. from a keratinocyte cDNA library. Pub polypeptides can be produced in transfected host cells. They can control the steady state level of cdc25 phosphatase, the degree of CDK kinase (e.g. cdc2) dephosphorylation and the steady state level of p53 (controlling the degree of cell cycle regulation of p53). They can regulate cell or tissue differentiation, or cell growth or proliferation by affecting other proteins, can be a specific (anti)agonist of wild-type CC protein function and may be used as immunogens to elicit a specific immune residue

XX Sequence 748 AA;

Query Match 99.8%; Score 4031; DB 2; Length 748;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MSNPGRRNGPVKRLITVLCANLVKDFPLPPFAKVVVDGSGQCHSTDTVKNLDPK 60
DB 1 MSNPGRRNGPVKRLITVLCANLVKDFPLPPFAKVVVDGSGQCHSTDTVKNLDPK 60
QY 61 WNOHVDLYIGKSDSYTISVWNNHKKIKKKOGAGFLGCVRLSNAINRLKDTGYORLDCKL 120
DB 61 WNOHVDLYIGKSDSYTISVWNNHKKIKKKOGAGFLGCVRLSNAINRLKDTGYORLDCKL 120
QY 121 GPNNDIVRGQIVVSLQSRDRIIGTGGVYVDCSRLFDNDLPDGEERRTASGRIOYLNHIT 180
DB 121 GPNNDIVRGQIVVSLQSRDRIIGTGGVYVDCSRLFDNDLPDGEERRTASGRIOYLNHIT 180
QY 181 RTQWERPTRPASESSSPGRPLSCFVDENPTISGTNATCGSSDPLARRVRSQHRN 240
DB 181 RTQWERPTRPASESSSPGRPLSCFVDENPTISGTNATCGSSDPLARRVRSQHRN 240
QY 241 YMSRTHLHTPPDLDEGEYQRTTQGGVYFLHTQTGVSFWHDPRPRDLNSINCEELGFLP 300
DB 241 YMSRTHLHTPPDLDEGEYQRTTQGGVYFLHTQTGVSFWHDPRPRDLNSINCEELGFLP 300
QY 301 PGWEIRNTATGRVYFVDDNNRTTQTPRISANHLVLNQNOLKDOQQQVYVSLCPDDT 360
DB 301 PGWEIRNTATGRVYFVDDNNRTTQTPRISANHLVLNQNOLKDOQQQVYVSLCPDDT 360
QY 361 ECLTVPRYKRDLYOKLILRQELSQOQPOAGHCRIEVSRREIPEESYROYMKRXPDLWK 420
DB 361 ECLTVPRYKRDLYOKLILRQELSQOQPOAGHCRIEVSRREIPEESYROYMKRXPDLWK 420

```

```

Db      361 ECLTPRYKRDVOKIKILRDELSCQOPAGHCRIEVSHEEIFFEBSYQVMKRPKDIWK 420
QY      421 RLMIKRGEGGLDYGVAREWILYLISHENLNFYGGFYQSDDIYTLQINPDSAVNPEHL 480
Db      421 RLMIKRGEGGLDYGVAREWILYLISHENLNFYGGFYQSDDIYTLQINPDSAVNPEHL 480
QY      481 SYFHFVGRIMGAVPHGHYIDGGFTLPFYKOLLGKSITLDMELVDPDLNLSVLTLEND 540
Db      481 SYFHFVGRIMGAVPHGHYIDGGFTLPFYKOLLGKSITLDMELVDPDLNLSVLTLEND 540
QY      541 TNGVLDHTECVHMAVGEIIOHEILKNGKSIPVNEENKSEYRLVYNNRFLGIAFOFLA 600
Db      541 TNGVLDHTECVHMAVGEIIOHEILKNGKSIPVNEENKSEYRLVYNNRFLGIAFOFLA 600
QY      601 LQKGFNEVLPQHLKTFDEKELELIIGLQKIDVNDKVNTRLKCTPDSDIVKMPKXAV 660
Db      601 LQKGFNEVLPQHLKTFDEKELELIIGLQKIDVNDKVNTRLKCTPDSDIVKMPKXAV 660
QY      661 EFPDEERRARLLOFTVGSSRVPLOGFKALQGAAGPRLFTIHQIDACTNNLPRKATCENRI 720
Db      661 EFPDEERRARLLOFTVGSSRVPLOGFKALQGAAGPRLFTIHQIDACTNNLPRKATCENRI 720
QY      721 DIPPESEYKLYEKLTALEETGFAVE 748
Db      721 DIPPESEYKLYEKLTALEETGFAVE 748

RESULT 3
AAU19610
ID AAU19610 standard; protein; 804 AA.
XX
AC AAU19610;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human diagnostic and therapeutic polypeptide (DITHP) #196.
XX
DE Human, receptor; diagnostic; therapeutic; gene therapy; vaccine;
XX
DE cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;
XX
DE acquired immune deficiency syndrome; AIDS; autoimmune disorder;
XX
DE respiratory disorder.
XX
OS Homo sapiens.
XX
PN WO200162927-A2.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US006059.
XX
XX
PR 24-FEB-2000; 2000US-0184693P.
PR 24-FEB-2000; 2000US-0184697P.
PR 24-FEB-2000; 2000US-0184698P.
PR 24-FEB-2000; 2000US-0184768P.
PR 24-FEB-2000; 2000US-0184769P.
PR 24-FEB-2000; 2000US-0184770P.
PR 24-FEB-2000; 2000US-0184771P.
PR 24-FEB-2000; 2000US-0184772P.
PR 24-FEB-2000; 2000US-0184773P.
PR 24-FEB-2000; 2000US-0184774P.
PR 24-FEB-2000; 2000US-0184776P.
PR 24-FEB-2000; 2000US-0184777P.
PR 24-FEB-2000; 2000US-0184779P.
PR 24-FEB-2000; 2000US-0184813P.
PR 24-FEB-2000; 2000US-0184837P.
PR 24-FEB-2000; 2000US-0184841P.
PR 24-FEB-2000; 2000US-0185213P.
PR 24-FEB-2000; 2000US-0185216P.
PR 12-MAY-2000; 2000US-0203785P.
PR 15-MAY-2000; 2000US-0204226P.
PR 16-MAY-2000; 2000US-0204525P.
PR 16-MAY-2000; 2000US-0204821P.
PR 16-MAY-2000; 2000US-0204908P.

```

```

PR 16-MAY-2000; 2000US-0205232P.
PR 17-MAY-2000; 2000US-0204815P.
PR 17-MAY-2000; 2000US-0204833P.
PR 17-MAY-2000; 2000US-0205212P.
PR 17-MAY-2000; 2000US-0205285P.
PR 17-MAY-2000; 2000US-0205286P.
PR 17-MAY-2000; 2000US-0205287P.
PR 17-MAY-2000; 2000US-0205323P.
PR 17-MAY-2000; 2000US-0205324P.
XX
XX
PA (INCY-1) INCYTE GENOMICS INC.
XX
PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC,
PI Chen A, D'sa SA, Amshery S, Dahl CR, Dam TC, Daniels SE, Dufour GE,
PI Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AL, Liu TF,
PI Roseberry AM, Rosen BH, Russo PD, Stockbrecher TK, Daffo A,
PI Wright RJ, Yap PE, Yu YJ, Bradley DU, Bratcher SR, Chen W,
PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S,
XX
XX
DR MPI, 2001-502867/55.
DR N-PEDB; AAS31181.
XX
PT polynucleotides encoding diagnostic and therapeutic proteins, e.g.
PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.
XX
PS Claim 27; Page 512-514; 522pp; English.

```

The invention relates to polynucleotides (I) encoding diagnostic and therapeutic (DITHP) polypeptides (II), which include e.g. enzymes, and proteins involved in growth and development and receptors. (I) and (II) may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate DITHP expression. For example, (I) and (II) may be used to treat disorders associated with decreased polypeptide expression by rectifying mutations or deletions in a patient's genome, that affect the activity of the DITHPs, by expressing inactive proteins or supplementing the patient's own production of them. (I) and (II) may be used to treat diseases, for example, cell proliferative disorder, Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma, leukaemia, autoimmune disorders, and respiratory disorders. Additionally, (I) may be used to produce the DITHPs, by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) and (II) may be used to produce the DITHPs, by inserting the nucleic acids into complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acids in samples, and therefore which patients may be in need of restorative therapy. (II) may also be used as antigens in the production of DITHP antibodies against DITHPs and in assays to identify modulators of DITHP expression and activity. The anti-DITHP antibodies and antagonists may also be used to down regulate expression and activity. The anti-DITHP antibodies may also be used as diagnostic agents for detecting the presence of DITHPs in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). AAU19415-AAU19625 represent human diagnostic and therapeutic (DITHP) polypeptides of the invention

Sequence 804 AA;

Query Match 99.7%; Score 4027; DB 4; Length 804;  
 Best Local Similarity 99.7%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;  
 Matches 746; Conservative 0;

```

QY 1 MSNPGRRNGPYKRLTYLCAKLVKDFRLPDPFAKVVDGSGQCHSTDTVKNLTDPK 60
Db 57 MSNPGRRNGPYKRLTYLCAKLVKDFRLPDPFAKVVDGSGQCHSTDTVKNLTDPK 116
QY 61 KNOHDLVYKGSDDSTISVWNRKTHKQAGFLGCVLLSLAIRLDTGQRLDCKL 120
Db 117 KNOHDLVYKGSDDSTISVWNRKTHKQAGFLGCVLLSLAIRLDTGQRLDCKL 176
QY 121 GPNDNDTVAGQIVSLQSRDRIGTGQVVDGSRLLPDPGMBERRRTASGRIOYLNHT 180
Db 177 GPNDNDTVAGQIVSLQSRDRIGTGQVVDGSRLLPDPGMBERRRTASGRIOYLNHT 236
QY 181 RTTWERPTTPASGEYSPPRLPSTCVNDTPISGNGATCGGSSPRLAERVRSQRRRN 240

```

```

Db      237 RTTWERPPTPRASSEYSGRPLSCFVDNTPISGTMGATCGSSDPRLAERRVRSQRHN 296
Qy      241 YKSRTHLHTPPDLPEGYQRTTQGOVYFLHTQGVSTWMDPRVPRDLSINCEELGP 300
Db      297 YMSRTHLHTPPDLPEGYEQRRTQGOVYFLHTQGVSTWMDPRVPRDLSINCEELGP 356
Qy      301 FGMELRNTATGVRVVDHNNRTTQFTDPRLSANHLVLRONOLKDQOQOQVSLCPDDT 360
Db      357 FGMELRNTATGVRVVDHNNRTTQFTDPRLSANHLVLRONOLKDQOQOQVSLCPDDT 416
Qy      361 ECLIVPRYKDLVQKLKILRQELSQOQFQACHCRIEVSRREIFESYRQVMMKPKDLMK 420
Db      417 ECLIVPRYKDLVQKLKILRQELSQOQFQACHCRIEVSRREIFESYRQVMMKPKDLMK 476
Qy      421 RLMTFRGEEGLDGGVAREWLYLISHMLNPPYGLPQYSRDDITVTQINPDSVNEHL 480
Db      477 RLMTFRGEEGLDGGVAREWLYLISHMLNPPYGLPQYSRDDITVTQINPDSVNEHL 536
Qy      481 SYFHFVGRIMGAVFHHGYIDGGFTLPFYKQLCKSKITLDDMELVDEPLHNSLWILEND 540
Db      537 SYFHFVGRIMGAVFHHGYIDGGFTLPFYKQLCKSKITLDDMELVDEPLHNSLWILEND 596
Qy      541 ITGVLDHTEFCVEHNAVGGIILQELKPNKSIPIVNEENKKEVRLYVWRFRLGIEAOFLA 600
Db      597 ITGVLDHTEFCVEHNAVGGIILQELKPNKSIPIVNEENKKEVRLYVWRFRLGIEAOFLA 656
Qy      601 LQKGENEYIPQHLKTFDEKELLELIGLGIQDVNDKVNTRLGCTPDNSIVKMFYKAV 660
Db      657 LQKGENEYIPQHLKTFDEKELLELIGLGIQDVNDKVNTRLGCTPDNSIVKMFYKAV 716
Qy      661 EFPDEERARLLQVYTGSRRVPLQGFALQGAAGPRFLTTHQIDACTNNLPKHAHTCENRI 720
Db      717 EFPDEERARLLQVYTGSRRVPLQGFALQGAAGPRFLTTHQIDACTNNLPKHAHTCENRI 776
Qy      721 DIPPESEYKLYEKLTLAIETCGFAVE 748
Db      777 DIPPESEYKLYEKLTLAIETCGFAVE 804

RESULT 4
AAE32724
ID      AAE32724 standard; protein; 722 AA.
XX
AC      AAE32724;
DT      24-MAR-2003 (first entry)
XX
DE      E3 ubiquitin ligase SMURF1 protein.
XX
KW      Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;
KW      poliomyelitis; HIV; measles; protein therapy; E3 ubiquitin ligase;
KW      enzyme.
XX
OS      unidentified.
XX
PN      MO200290549-A2.
PD      14-NOV-2002.
XX
PF      12-MAR-2002; 2002W0-IB002106.
XX
PR      12-MAR-2001; 2001US-0275224P.
PR      31-JUL-2001; 2001US-0308958P.
PR      07-DEC-2001; 2001US-0340170P.
XX
PA      (PROT-) PROTEOLOGICS LTD.
XX
PI      Greener T, Moskowitz H, Reiss Y, Alroy I;
PI      WPI; 2003-111976/10.
DR      N-PSDB; AAD50458.
XX
PT      New protein complex comprising HECT-RCC1, viral maturation scaffolding

```

protein (WMS), and/or HIV gag protein, useful for treating viral infections, such lymphosarcoma, HIV, hepatitis, poliomyelitis, measles, or Ebola.

Disclosure; Fig 14; 150pp; English.

The invention relates to a method for modulation of viral maturation. The invention also provides an isolated protein complex comprising a HECT-RCC1 selected from HECT-RW, HECT-RCC1, Gag protein, Gag late domain, P13, actin, myosin, Hsp60, Hsp90, STAM1, STAM2, VHS-VIM, CC, GPase, E2 enzyme, tag101, cullin, HECT1, HECT2, HECT3, Nedd4-like protein or clathrin. The complexes, proteins, antibodies and methods are useful for treating viral infections, such as lymphosarcoma, human CC immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola and for inhibiting budding in a subject. They are also useful in diagnostic assays for determining whether a cell is infected with a virus and for characterizing the nature, progression and/or infectivity of the infection. The invention is also useful in protein therapy. The present sequence is E3 ubiquitin ligase SMURF1 protein used to illustrate the method of the invention

Sequence 722 AA:

Query Match 74.3%; Score 3001; DB 6; Length 722;  
Best Local Similarity 74.6%; Pred. No. 8e-273;  
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

```

Qy      12 VKLRLTVCAKLVKKDFFRLPDPRAKYVVGSGGCHSTDPVKNLDPKMNQHDLYIGK 71
Db      4 IKRLTLVCAKLVKKDFFRLPDPRAKYVVGSGGCHSTDPVKNLDPKMNQHDLYIGK 63
Qy      72 SDSVTISVNNHKKIKHKKOAGAGLGCVRLLSNAINLKDQYQRLDCLGPNDDTVRGQ 131
Db      64 TDSITISVNNHKKIKHKKOAGAGLGCVRLLSNAINSLKDTQYQRLDCLGPNDDTVRGQ 123
Qy      132 IVYSIQSRDRIGTGGQVVDCSRFPNDLPDQWBERRTASGIIQVNLHTTRTQWBRFRP 191
Db      124 IVYSIQSRDRIGTGGQVVDCSRFPNDLPDQWBERRTASGIIQVNLHTTRTQWBRFRP 154
Qy      192 ASEYSGRPLSCFVDNTPISGTMGATCG-----QSDPRLAERRVRSQRHNYM 242
Db      155 --EDSGRPLSCFVDNTPISGTMGATCG-----QSDPRLAERRVRSQRHNYM 212
Qy      243 ---SRTHLHTPPDLPEGYEQRRTQGOVYFLHTQGVSTWMDPRVPRDLSINCEELGP 298
Db      213 QTPQNRPHGHOSPELPEGYEQRRTQGOVYFLHTQGVSTWMDPRVPRDLSINCEELGP 272
Qy      299 LPPGWEIRNTATGVRVVDHNNRTTQFTDPRLSANHLVLRONOLKDQOQOQV----S 354
Db      273 LPPGWEIRNTATGVRVVDHNNRTTQFTDPRLSANHLVLRONOLKDQOQOQV----S 328
Qy      355 LCPDDETLVPRYKDLVQKLKILRQELSQOQFQACHCRIEVSRREIFESYRQVMMK 414
Db      329 L---EDELPQRYERDLVQKLKILRQELSQOQFQACHCRIEVSRREIFESYRQVMMK 385
Qy      415 PKDLMKRLMTFRGEEGLDGGVAREWLYLISHMLNPPYGLPQYSRDDITVTQINPDSV 474
Db      386 PKDLMKRLMTFRGEEGLDGGVAREWLYLISHMLNPPYGLPQYSRDDITVTQINPDSV 445
Qy      475 VNEBELSYFHFVGRIMGAVFHHGYIDGGFTLPFYKQLCKSKITLDDMELVDEPLHNSLV 534
Db      446 INPDLSYFHFVGRIMGAVFHHGYIDGGFTLPFYKQLCKSKITLDDMELVDEPLHNSLV 505
Qy      535 WLENDITGVLDHTEFCVEHNAVGGIILQELKPNKSIPIVNEENKKEVRLYVWRFRLG 594
Db      506 WLENDITGVLDHTEFCVEHNAVGGIILQELKPNKSIPIVNEENKKEVRLYVWRFRLG 565
Qy      595 EAOFLAQKGENEYIPQHLKTFDEKELLELIGLGIQDVNDKVNTRLGCTPDNSIVK 654
Db      566 EAOFLAQKGENEYIPQHLKTFDEKELLELIGLGIQDVNDKVNTRLGCTPDNSIVK 625
Qy      655 WFKAVFEFDEERARLLQVYTGSRRVPLQGFALQGAAGPRFLTTHQIDACTNNLP 711
Db      625 WFKAVFEFDEERARLLQVYTGSRRVPLQGFALQGAAGPRFLTTHQIDACTNNLP 688

```

[illegible][illegible]

20-JUN-2000; 2000US-00598075.  
 PR 19-JUL-2000; 2000US-00620325.  
 PR 01-SEP-2000; 2000US-00654936.  
 PR 15-SEP-2000; 2000US-00653561.  
 PR 20-OCT-2000; 2000US-00653325.  
 PR 30-NOV-2000; 2000US-00728422.  
 XX  
 PA (HYSE-) HYSEQ INC.

XX Tang YF, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Gao Y;  
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
 XX WPI: 2001-476283/51.  
 DR N-PDB; AAKS2994.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 PT in diagnosis and gene therapy.

XX Claim 20; Page 362-363; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAW76323-AAW80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Record for SEQ ID NO 2110 (AAK55581), 2111  
 CC (AAK52882) and 3666 (AAW80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 XX  
 XX Sequence 722 AA;

XX Query Match 74.1%; Score 2992; DB 4; Length 722;

XX Best Local Similarity 74.5%; Pred. No. 5.6e-272;

XX Matches 564; Conservative 63; Mismatches 72; Indels 58; Gaps 8;

QY 12 VKRLTVLCANLVKQEFFRLDPFAVAVVVDGSGQCHSTDTVKNLTDPKKNHXYDTYTK 71  
 DB 4 IKRLTVLCANLVKQEFFRLDPFAVAVVVDGSGQCHSTDTVKNLTDPKKNHXYDTYTK 63  
 QY 72 SDSVTISVNMHKKTHKQAGAGFAGCVLLSNALNRLKDTGYQRLDCKLGNNDVTRGQ 131  
 DB 64 TDSITISVNMHKKTHKQAGAGFAGCVLLSNALNRLKDTGYQRLDCKLGNNDVTRGQ 123  
 QY 132 IVVSLQSRDRIGTGGVVDGSRLLFDNDLPDGEERRTASGRIGYLNHITRTQWERPTR 191  
 DB 124 IVVSLQSRDRIGTGGVVDGSRLLFDNDLPDGEERRTASGRIGYLNHITRTQWERPTR 154  
 QY 192 ASESSSGRPLSCVVDENTPISGNTGATCG-----GSPRLAERRVSQRHRNTY 242  
 DB 155 --EDSGGRPLSCVMEBPAPYDTSTGAAGGNCRFVESPSCQGRACQRLRNPDRGSI 212  
 QY 243 ---SRTHLATPPLPEGYEORTTQGGVYFLHTQTVSWPDPVPRDLSNNGEILP 298  
 DB 213 QTPQNRHGHGOSPELPEGYEORTTQGGVYFLHTQTVSWPDPVPRDLSNNGEILP 272  
 QY 299 LPPGWEIRNTATGAVYVVDHNNRTTQFTDRLSANTLHLVNRNOQKDOOQGVV---S 354  
 DB 273 LPPGWEIRNTATGAVYVVDHNNRTTQFTDRLSANTLHLVNRNOQKDOOQGVV---S 328  
 QY 355 LCPDTECLVPRYKRLVQKTLIRQELSQOQPOGHCHIEVSRREIFESRQVMKR 414  
 DB 329 L---EDELPAQRYERLVOKLVKRLHSLQOPQGHCHIEVSRREIFESRQVMKR 385  
 QY 415 PKDLMKRLMKFRGEGLDYGVAREWLYLLSHEMLNPPYGLFQYARDIYTLQINPDA 474  
 DB 386 PKDLMKRLMKFRGEGLDYGVAREWLYLLSHEMLNPPYGLFQYARDIYTLQINPDA 445

QY 475 VNEPHSYEPFVGRIGMAVFRGHYIDGAGFTLPFYKQLLGSITLDDMELVDPDLNSLV 534  
 DB 446 INPDHLSYEPFVGRIGMAVFRGHYIDGAGFTLPFYKQLLGSITLDDMELVDPDLNSLV 505  
 QY 535 WLENDITGVLDATPCVENAVGEIIOHCLKNGKSIPIVNEKKYVLYLVNMFELRGI 594  
 DB 506 WLENDITGVLDATPCVENAVGEIIOHCLKNGKSIPIVNEKKYVLYLVNMFELRGI 565  
 QY 595 EAQFLALQGFNEVITQHLKTPDEKELELITCGSLKIDVNDKRNTRKXCTPPSNIVK 654  
 DB 566 EAQFLALQGFNEVITQHLKTPDEKELELITCGSLKIDVNDKRNTRKXCTPPSNIVK 625  
 QY 655 WFWKAVEFPDEBRRAALLQFTVSSRPVQGFKAQ---GAAPRLFTHOIDACNNLP 711  
 DB 626 WFWKAVEFPDEBRRAALLQFTVSSRPVQGFKAQ---GAAPRLFTHOIDACNNLP 685  
 QY 712 KATCFNRIDIPPYSEYKYLITALETCGFAVE 748  
 DB 686 KATCFNRIDIPPYSEYKYLITALETCGFAVE 722

# RESULT 7

AAU87301 standard; protein, 514 AA.

AAU87301;  
 05-JUN-2002 (first entry)

Novel central nervous system protein #211.

Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
 hyperproliferative disorder; neoplasm; cardiovascular disorder;  
 cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;  
 nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
 acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
 adenocarcinoma; reproductive system disorder; testicular feminisation;  
 endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
 respiratory disorder; renal disorder; kidney failure; blood disorder;  
 myocardial infarction; wound healing; cell proliferation; skin aging;  
 food additive; food preservative; gene therapy.

Homo sapiens.

WO200155318-A2.

02-AUG-2001.

17-JAN-2001; 2001WO-US001332.

PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0214886P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 07-JUL-2000; 2000US-0216860P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217486P.  
 PR 14-JUL-2000; 2000US-0218290P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224518P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0228344P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0233080P.  
PR 08-SEP-2000; 2000US-0233081P.  
PR 12-SEP-2000; 2000US-0233968P.  
PR 14-SEP-2000; 2000US-0233979P.  
PR 14-SEP-2000; 2000US-0233989P.  
PR 14-SEP-2000; 2000US-0234001P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234222P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0244612P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249255P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 06-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-581633/65.  
XX N-PSDB; ABK43631.  
DR  
XX  
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,  
PT treating or ameliorating medical conditions and used as food additives or  
XX preservatives.  
XX  
PS Claim 9; SEQ ID NO 819; 837pp; English.  
XX  
XX  
CC The invention describes an isolated nucleic acid molecule (I) encoding a  
CC novel central nervous system protein. (I) and polypeptides (II) encoded  
CC by (I) are used to treat a medical condition and in diagnosis of a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia,  
CC angiosgenesis, nervous system disorders e.g. Alzheimer's disease and  
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses  
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,  
CC adenocarcinomas and irritable bowel syndrome, reproductive system  
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes  
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.  
CC leukaemia, disorders involving neovascularisation e.g. malignancies,  
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
CC acute kidney failure and blood related disorders e.g. myocardial  
CC infarction. The polypeptides can also be used to aid wound healing and  
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to

CC maintain organs before transplantation, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 68.7%; Score 2775; DB 4; Length 514;  
Best Local Similarity 99.8%; Pred. No. 8.9e-252;  
Matches 513; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 235 SCRRNNYSRTHLHTPPDLPEGYEQRTTQCGVYFLHTQTGVSTWMDPRVRLDLSNINCE 294
DB 1 SRRHNNYSRTHLHTPPDLPEGYEQRTTQCGVYFLHTQTGVSTWMDPRVRLDLSNINCE 60
QY 295 ELGRLPGMEINTATGVRVYVDHNNRTQFTDPLSANLHLYNRQQLDQOQOQVVS 354
DB 61 ELGRLPGMEINTATGVRVYVDHNNRTQFTDPLSANLHLYNRQQLDQOQOQVVS 120
QY 355 LCPDDTECTVPRYKRDVQKILROELSOQOPAGHCRIEVSREIEFEESYQVMKMR 414
DB 121 LCPDDTECTVPRYKRDVQKILROELSOQOPAGHCRIEVSREIEFEESYQVMKMR 180
QY 415 PKDLKRLMKRREGGLDYGVARFEMLYLSHENTANYGLFQYSRDIYTLQINPDSA 474
DB 181 PKDLKRLMKRREGGLDYGVARFEMLYLSHENTANYGLFQYSRDIYTLQINPDSA 240
QY 475 VNPEHLSEYFHFVGRMGAVFEGHYIDGGFTLPFYKQLLGSITLDEVELVDPDLNSLV 534
DB 241 VNPEHLSEYFHFVGRMGAVFEGHYIDGGFTLPFYKQLLGSITLDEVELVDPDLNSLV 300
QY 535 WLENDITGVLDHTCEVHNAYAGETIQHELKPNKSIPIVNEENKKEYRLVYVNRFLRGI 594
DB 301 WLENDITGVLDHTCEVHNAYAGETIQHELKPNKSIPIVNEENKKEYRLVYVNRFLRGI 360
QY 595 EAQFLALOKGFENEVPOHLKLTDFEKELELLICGKIDVNDWVKNRLKHCPTDSNIYK 654
DB 361 EAQFLALOKGFENEVPOHLKLTDFEKELELLICGKIDVNDWVKNRLKHCPTDSNIYK 420
QY 655 WFWKAVEFPDEERRARLLQVYTGSSRVPLQGFKALOGAAGRLLTTHQIDACTNNLPKX 714
DB 421 WFWKAVEFPDEERRARLLQVYTGSSRVPLQGFKALOGAAGRLLTTHQIDACTNNLPKX 480
QY 715 TCENRIDIPPESEYEKLEKLLTAIEETCGFAVE 748
DB 481 TCENRIDIPPESEYEKLEKLLTAIEETCGFAVE 514

RESULT 8
ABB61120
ID ABB61120 standard; protein; 1035 AA.
AC ABB61120;
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 10152.
DX Drosophila: developmental biology; cell signalling; insecticide;
KW pharmaceutical.
OS Drosophila melanogaster.
XX WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009221.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKB) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;

```

XX WPI; 2001-656860/75.  
DR N-PSDB; ABL05223.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.

PS Disclosure; SEQ ID NO 10152; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-  
CC AB578072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1035 AA;

Query Match 53.8%; Score 2173; DB 4; Length 1035;  
Best Local Similarity 43.2%; Pred. No. 1.4e-194;  
Matches 456; Conservative 104; Mismatches 154; Indels 342; Gaps 16;

```

QY 7 RRRGPKALRLTYLCAKLVKDFRLPDPFAVYVDGSGQCHSTDTYKNTLDPKNOHYD 66
DB 8 RRRGTHKVRITTLICANNLRKDLFRLPDPFAVYVDGSGQCHSTDTYKNTLDPKNAHYD 67
QY 67 LVYKSDSVTIIWVNNHKIKHKKGAGFLGCYLLNAINRLKDYQRLDCKLGPNNND 126
DB 68 LFLGIGDATTIVYNNQRKHK--GSGFLGCVRIAPFNQSLKAGAFGLDLCKLSPDDE 125
QY 127 TYRQIVLSLQSRDR-----IGTGGVVDGSR--LFDNDLPDGEERTASGRIOY 175
DB 126 LVYRQIILSLKSDGSSGNPLAIVGSDVYRGPEDSDSESLPEBGEERTDNGRYVY 185
QY 176 LNHITPTQWERTPR----- 191
DB 186 VNHAKISTOMDPRQPGVYVSSSHATSPOQRANTHANGSGDROAPAGPRTSTCTNNMNG 245
QY 192 -----ASEYSSPGRPLSCFV----- 206
DB 246 HRSRLSVTASDERHSTEILSSVCKENTSPFTPVASATTPGKKTSSNSSASAGRTLEQ 305
QY 207 -----DENTPISGT-----N 216
DB 306 RPTNEPAITPSTTSASVFLHSNDNHVKTPKHQTNGHAPPESTPTSPYQOQVYVNGAQN 365
QY 217 GAT-----CGQSSDPRLAER----- 231
DB 366 GSTSGNGSGQAAPQPSANGMTQEDAAATTSSTTSPRHSQSPPTNISPASVYPSA 425
QY 232 -----RYASQHRNYMGRTHLHTPP----- 251
DB 426 NGNVHSPNANSTPAGSGGGSRYTATPGORSORSSQCGEESTRRSSRGTRNGTS 485
QY 252 -----DLEPGYEQRTTQCGVYFLHTQTGVSTWMDPR 294
DB 486 GGGGGGGGQRYASAAIAAQAARPPUDLPFGYEMRTTQGGVYFHHIPGVSTWMDPR 545
QY 284 VPRDL--SNINCEIIGPLPGMEIRNTATGVRVYVDHNNRTQFTDPLSANLHLYNRQ 341
DB 546 IPRDFDTQHLTLDAIGPLPGMEIRNTATGVRVYVDHNNRTQFTDPLSANLHLYNRQ 605
QY 342 N-----QLNDQOQOQVVS-----CPDTE 361
DB 606 TVPPTSAANAGTPAPSPATPAPSAAAVPPQATPASNATPTTLTTTNPHRRLVPLDLP 665
QY 362 C-----TVPRYKRDVQKILROELSOQOPAGHCRIEVSREIEFEESYQVMKMRPK 416

```

Db 666 GLEGGADLPKYRDLVGLRLALRTELQTMOPQSGHCRLEVSARNEIFEESYLLIMKRAK 725  
 Qy 417 DLAKRMIFKRGEGLDYGGVAREWYLISHMLNPPYGLFOYSRDIYTLQINPDASN 476  
 Db 726 DMRKRLMVFKEGEGLDYGGVAREWYLISHMLNPPYGLFOYSRDIYTLQINPDASN 785  
 Qy 477 PEHLISYFHEVGRIMGNVAFHGHYIDGGFTLLPFYKQLGKSTLDDMELVDPDLANSIWT 536  
 Db 786 PDHLISYFHEVGRIMGNVAFHGHYIDGGFTLLPFYKQLGKSTLDDMELVDPDLANSIWT 819  
 Qy 537 LENDIGVADHPCVEHNNAVGEIICHELKPNKSPVNEENKKEYRLVNMNRPGLGIA 596  
 Db 820 LBNISIGLLESTPVSNNSGALVHLELPGGASIPTEENKREYKLVNRRFNGIQQ 879  
 Qy 597 QFLALQKGFNEVLPQHLKTFDEKELELLICGLKIDVNDMKNVTELKCTPDNSITVWF 656  
 Db 880 QFLALQKGFCELLPSHLRPFDERELVIGGISIDVNDMKNVTELKCTPDNSITVWF 939  
 Qy 657 WKAVEFDEERRARLLIQFTGSSRVLQGFKAQ---GAAGPRLFTIH-QIDACTNNLPK 712  
 Db 940 WQVVESSSEMRARLLIQFTGSSRVLQGFKAQ---GAAGPRLFTIH-QIDACTNNLPK 999  
 Qy 713 AHTCFNRIDIPPESEYEKLYEKLTAIEETCGFAVE 748  
 Db 1000 AHTCFNRIDIPPESEYEKLYEKLTAIEETCGFAVE 1035  
 RESULT 9  
 ID ADC64238 standard; protein; 335 AA.  
 AC ADC64238;  
 XX 18-DEC-2003 (first entry)  
 DT Human SMURF2\_v2.  
 DE  
 XX Hepatotropic; vaccine; protein-protein interaction;  
 KM Transforming Growth Factor beta; TGF beta; hepatitis;  
 KM Selected Interacting Domain; SID; bait; human; SMURF2\_v2.  
 OS Homo sapiens.  
 XX  
 PN W02003045990-A2.  
 PD 05-JUN-2003.  
 PF 26-NOV-2002; 2002WC-EP013866.  
 XX  
 PR 26-NOV-2001; 2001US-0333348P.  
 PR 31-MAY-2002; 2002US-0384537P.  
 PR 30-OCT-2002; 2002US-0422471P.  
 PA (HYBR-) HYBRIGENICS.  
 XX  
 PI Legrain P, Gauthier J, Colland F, Jacq X;  
 DR WPI; 2003-505185/47.  
 DR N-PSDB; ADC64442.  
 XX  
 PT New complex between two interacting proteins, useful for screening  
 PT molecules that inhibit transforming growth factor beta (TGF beta) or TGF  
 PT beta super-family of cytokines pathway for diagnosing or treating TGF  
 PT beta diseases or disorders.  
 XX  
 PS Disclosure; SEQ ID NO 22; 148bp; English.  
 CC The present invention relates to protein-protein interactions and  
 CC complexes involved in Transforming Growth Factor (TGF) beta disorders  
 CC and/or diseases. The complex between two interacting proteins is useful  
 CC for screening molecules that inhibit TGF beta for diagnosing or treating  
 CC diseases or disorders involving TGF beta e.g., hepatitis. To illustrate  
 CC the invention, selected interacting domains (SID) of proteins and their

CC coding sequences were isolated (ADC64238-ADC64318). Sequences ADC64318-  
 CC ADC64242 are bait sequences used to isolate the SIDs of the invention.  
 CC  
 SQ Sequence 335 AA;  
 Query March 44.8%; Score 1809; DB 7; Length 335;  
 Best Local Similarity 99.1%; Pred. No. 4,3e-161;  
 Matches 332; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 MSNPGRRNGFVKRLRYLCAKLVKDFRLPDPFAKVVVDSGSGCHSTDTVKNITDPK 60  
 Db 1 MSNPGRRNGFVKRLRYLCAKLVKDFRLPDPFAKVVVDSGSGCHSTDTVKNITDPK 60  
 Qy 61 WNGHYDLYIGSDSVTTSVWNNKKIKKQAGFLGCVRLSNMINKDGYQSLDCKL 120  
 Db 61 WNGHYDLYIGSDSVTTSVWNNKKIKKQAGFLGCVRLSNMINKDGYQSLDCKL 120  
 Qy 121 GPNDNDIVRQIIVSLQSRDRICTGGQVVDCSRLEFNDLPDGMERRTAGRIQYLNHIT 180  
 Db 121 GPNDNDIVRQIIVSLQSRDRICTGGQVVDCSRLEFNDLPDGMERRTAGRIQYLNHIT 180  
 Qy 181 RTTQWERPTPASPISGSPRPLSCFVDENPTISGTNGATCGQSSDPRLARVRVRSQSHRN 240  
 Db 181 RTTQWERPTPASPISGSPRPLSCFVDENPTISGTNGATCGQSSDPRLARVRVRSQSHRN 240  
 Qy 241 YMSRTHLHTPPDLPEGYQRCTTQGGQVYFLHTGTGVTWHDPRVPRDLNSINCELGPLP 300  
 Db 241 YMSRTHLHTPPDLPEGYQRCTTQGGQVYFLHTGTGVTWHDPRVPRDLNSINCELGPLP 300  
 Qy 301 PGWEIRNTATGRVYFVDANNRTQTFTDPRLSANLH 335  
 Db 301 PGWEIRNTATGRVYFVDANNRTQTFTDPRLSANLH 335  
 RESULT 10  
 ID AAW13385 standard; protein; 766 AA.  
 AC AAW13385;  
 XX 10-JUL-1997 (first entry)  
 DT Human protein ubiquitin ligase pub2.  
 DE  
 XX Human protein ubiquitin ligase pub2.  
 KM Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase; p53;  
 KM cell cycle; transgenic animal.  
 OS Homo sapiens.  
 XX  
 PN W09712962-A1.  
 PD 10-APR-1997.  
 PF 04-OCT-1996; 96WO-US015930.  
 XX  
 PR 04-OCT-1995; 95US-00539205.  
 PA (COLD-) COLD SPRING HARBOR LAB.  
 XX  
 PI Beach D, Caligiuri M, Nefsky B;  
 DR WPI; 1997-226206/20.  
 DR N-PSDB; AAT47041.  
 XX  
 PT Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved in  
 PT cdc25 phosphatase and p53 ubiquitination, and regulate cell growth and  
 PT proliferation.  
 XX  
 PS Claim 1; Page 80-84; 108bp; English.  
 CC Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and pub3  
 CC (AAW13386) are homologues of fission yeast pub1 (AAW13387) and were  
 CC identified from cDNA clones (AAT47040-42) obt'd. e.g. from a keratinocyte

CC cDNA library. Pub polypeptides can be produced in transfected host cells.  
 CC They can control the steady state level of cdc25 phosphatase, the degree  
 CC of CDK kinase (e.g. cdc2) phosphorylation and the steady state level of  
 CC p53 (controlling the degree of cell cycle regulation of p53). They can  
 CC regulate cell or tissue differentiation, or cell growth or proliferation  
 CC by affecting other proteins, can be a specific (ant)agonist of wild-type  
 CC protein function and may be used as immunogens to elicit a specific  
 CC immune residue  
 CC XX

SQ Sequence 766 AA;

Query Match 38.3%; Score 1545; DB 2; Length 766;  
 Best Local Similarity 40.8%; Pred. No. 1,2e-135;  
 Matches 336; Conservative 121; Mismatches 234; Indels 132; Gaps 18;

QY 1 MSNPRRNGPVKRLTLCAKLVKDPFRLPPPAKVVDGSGCHSDTVANTDPK 60  
 DB 1 MSNSAQR---RLRVIIAADGLYKRDVFRPPFAVLTVDGE-QHTTAAIKKTLNIPY 55  
 QY 61 WNOHYDLYIGKSDSTISVNNHKKIHKQAGFGCVALLSNANRLKDGQRL-DLC 118  
 DB 56 WNETFEVAVTNSITAIQVFPQKKF-KKKGGFGLVNLKRGVDLDAIGDEMLIDDK 114  
 QY 119 KLGPNDNTVVGQIVLSIQ----- 137  
 DB 115 KL--NENTVVGKTIINLSTIAQLTLQVSSAASAGARTSITNDPQSSKSSVSHPA 172  
 QY 138 -----SRDRIGGCGVVDCSRLEFDN-----DLPGMERERRASGRIDYINHTPTQ 184  
 DB 173 SSRAGSPTRDVAAPASAPSSSEPTFSSGEDOYGLPFGMECTNLCRTYVVDNHTSIT 232  
 QY 185 WERTP-----RPASEVSSPGRPLSCFVDENTPISGTNGATCGGSDPRLAERRVSRH 238  
 DB 233 WIRPLSSVAGAAAEHLSSAS--SANTVEGVCSSSNAA-----RRTASAVLT-- 279  
 QY 239 RNVSRTHLHTPPLPESGEORTQCGQVFLHTQGTGVTMHPDRVRLDNLN----- 292  
 DB 280 ---SNATTAGGELPGEBOEYTPBEGSPYFVDNHTTTTWDPRQOYIRSYGGENNAT 335  
 QY 293 -----CEELGFLPGEWEIRNTATGRVYFVDNNTTQFTPTRLSANLHLNQNOLKDQ 347  
 DB 336 IQQGVPSQLGFLPFGMEWRLNLTARVYFVDNHTKTMTWDPRLPSSL-----DQ 384  
 QY 348 QQQQVSLCPDPTCCLVPRYKRLVOKLILRELSSQ--QPOAGCRLEVSREELFEE 405  
 DB 385 ---NVPQYKRDPRKRLIYF--LSQPALHPLPGQCHIKVARNHIFED 425  
 QY 406 SYRQWTKRPKDLMKRLMIKFRGEGGLDYGVAEMLYLSHEMLNPLYGLFOYSRDIT 465  
 DB 426 SYAIMQASATDLKRLMIKFDGEDGLDYGLSREYFLLSHENKPNFYCLFEYSVDNY 485  
 QY 466 TLQINPDSAVNPEHLSTFHFVGRIMGAVRHGHYIDGGFTLLPYKQLGKSTLDMELV 525  
 DB 486 TLQINPSGINPEHLNPFKEIGRVIQGLAFHRRFVDAFFVVSFYKMLIQKVTLDMEGM 545  
 QY 526 DPDLNSLWTLNDINGVLDHTECVENHAYGELIQHELMKNKSPVNEENKKEVRLY 585  
 DB 546 DAERYRLWLDNDIDGVDLTTSVDCNCFGEVVTIDLPKNGRNEIYEENKREYVDLV 605  
 QY 586 VNWFFLGLIEAQLALQGFENEVIPOHLLTFPEKELELLICGLKIDVNMKVNTRLKH 645  
 DB 606 TVM-IQKRIEQQFVAFHEGFESELIPOELINVPDERELLLIGLISEIDMEDMKKHDDYS 664  
 QY 646 CTPEPSNVKMFVKAVERFEDEERRARILQVTGSSRVVLOQFKLQAGARLLTHIQIDA 705  
 DB 665 YSENDQILIKFWELMDWMSNEKSRLLQFTGTTSRLPVNFGKOLQSDGPRKFTIKAGE 724  
 QY 706 CTNNLPKATCFNRID-PPYSEYEKLEYKLTALIEETGFAVE 748  
 DB 725 -PNKLPRKATCFNRIDLPPYTSKDDHKLSTIABETIGRGE 766

RESULT 11

ABP73459  
 ID ABP73459 standard; protein, 832 AA.  
 AC ABP73459;  
 DT 30-JAN-2003 (first entry)  
 XX  
 DE Candida albicans essential protein SEQ ID NO 7296.  
 XX  
 XX Fungus; yeast; tetracycline; promoter; GRACE strain; biosynthesis;  
 KW signal transduction; DNA replication; cell division; growth;  
 KM proliferation; Candida albicans; fungicide; antifungal.  
 XX  
 OS Candida albicans.

XX  
 XX MO200253728-A2.  
 XX  
 XX 11-JUL-2002.  
 XX  
 XX 26-DEC-2001; 2001WO-US049466.  
 XX  
 XX 29-DEC-2000; 2000US-0259128P.  
 XX 20-FEB-2001; 2001US-00792024.  
 XX 22-AUG-2001; 2001US-0314050P.  
 XX  
 XX (ELIT-) ELITRA PHARM INC.  
 XX  
 XX Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;  
 XX WPI; 2002-566694/60.  
 XX N-PSDB; ABZ32009.  
 XX  
 XX Constructing strains for identifying gene products as effective targets  
 XX for therapeutic intervention, by inactivating in the strain one allele of  
 XX a gene and placing other allele of the gene under conditional expression.  
 XX  
 XX Claim 4; SEQ ID NO 7296; 167bp + Sequence Listing; English.  
 XX  
 XX The invention relates to constructing (M1) a strain of diploid fungal  
 XX cells in which both alleles of a gene are modified, comprising modifying  
 XX one allele by insertion or replacement by a cassette having an  
 XX expressible selectable marker and modifying other allele by  
 XX recombination, of a promoter replacement fragment with a heterologous  
 XX promoter, so that expression of the second allele is regulated by the  
 XX promoter. (M1) is useful for constructing a strain of diploid fungal  
 XX cells in which both alleles of a gene are modified. The diploid fungal  
 XX cells having both alleles modified are useful for identifying a gene that  
 XX is essential to the survival or growth of a fungus, a gene that  
 XX contributes to the virulence and/or pathogenicity of a fungus, a gene  
 XX that contributes to the resistance of a diploid fungus to an antifungal  
 XX agent, an antifungal agent that inhibits the growth of a diploid fungus  
 XX and for identifying a therapeutic agent for treatment of a mammalian  
 XX disease. (M1) is useful for identifying a compound which modulates the  
 XX activity of a gene product, preferably enzymatic activity, carbon  
 XX compound catabolism, signal transduction, transporter, transcriptional,  
 XX translational, signal transduction, DNA replication and cell division  
 XX activity. The method is useful for identifying a compound having the  
 XX ability to inhibit growth or proliferation of C. albicans cells and for  
 XX treating infection by C. albicans. The present sequence is that of an  
 XX essential Candida albicans protein used in the method of the invention.  
 XX Note: The sequence data for this patent is not represented in the printed  
 XX specification but is based on sequence information supplied to Derwent by  
 XX the European Patent Office  
 XX  
 SQ Sequence 832 AA;

Query Match 37.9%; Score 1530.5; DB 5; Length 832;  
 Best Local Similarity 38.4%; Pred. No. 3.2e-134;  
 Matches 338; Conservative 120; Mismatches 228; Indels 195; Gaps 21;

QY 6 RRRGVPVRLTLVCAKLVKDPFRLPPPAKVVDGSGCHSDTVKNTLDPKKNQHY 65  
 DB 9 RSNNTTINVKVVAESLYKRDVFRQDPFAVLTVDS-QYKTTIRAKTLNPNWNETF 67

```

QY 66 DLYIGKSDSV--TISVNMNKKIHKKGAGFLGCVR-----LISNAINRLKDTGYORLDLCK 119
DB 68 N-FAKEDSILVIOVFQDKKF-KKDKQGLGIVNVRIGVDLISLNSSEITIR--DLKK 123
QY 120 LGPNDNDTVRGOIVVSLQSRDRIGTGGVVD----- 150
DB 124 --SNENLAVSGKIIYVI-SHNRNSGSGVTTATRTGTGASSNNIATITGVNLRIGSA 180
QY 151 -----CSRLFDN--- 157
DB 181 TTTNSTAQASSDATVANGSGPTSSLPPIGQHPETLAPDGAAGAASRQYSSFED 240
QY 158 --DLPDGEERRRATASGRIOYLNHTTTTOWERPTRPASE----- 194
DB 241 QYGRLLPQGWERTDNGRITVVDHNSRTTQWRPALQSETERSGQCRQSETEARRQHRG 300
QY 155 YSSPGR-----PLSCFVDENTPISGTNGATCGOSS-----DRLAERARVRSQHRN 240
DB 301 RTLPEGSVSL-----PTGSGNSTISGNTVNVNAGANTPVPAAVSMASGATTS 352
QY 241 YMSRTHLHTPDLPEGYDQRTTQGGVYFLHTQTGVSTWHDPRVPRDLNIN----- 292
DB 353 GLG-----ELPFGWEQRFTTEGRPYFVDHNRRTTWTVPFRQYIRTEGPTWTIQG 404
QY 293 -CEEIGLPPGWEINTATNGVYFVDHNRRTTQFTDPLSANLHLVLRNQNLKDOQQQ 351
DB 405 PVSQGLPSPGWEMLNTARVYFVDHNTKTTWDDPLPSL-----DQ--- 449
QY 352 VVSLCPDDTECLTVPRYKRDVLQKILROELSQQQPQ---AGCRIEVSREIEPESSY 407
DB 450 -----NVQYKDEPRFKATYR-----SQPALRILPQCHIKVRDRHIFDSY 492
QY 408 RQVMKMRPDLMLKMLTKFRGEGLDGCVARERMLXLSHEMNPPYGLQYSRDLYTL 467
DB 493 QIMKQTPEDLKKRLMKIFDGBEGLDVGVSREFPFLSHOMEPFCLFEYSHDNYTL 552
QY 468 QINPDSAVPEHLSTFHFVGRIMGAVNHGHIYDGGFTLPYKQLLSKITLDDMELVDP 527
DB 553 QINPDSGINPEHLNFKFGRVGLGVHRRFLAPFGALYKMLKHYVLQDMGVDA 612
QY 528 DLHNSLVMTLNDINGVLDHTRPCVEMHMXGELIQHEKPKNKSIPVNEKKEVRYLYN 587
DB 613 EYRRLKMLDNDIGLDLFTSAEESFGELIVEVDLPGGRDIEVTEENKHEVEVLITE 672
QY 588 WRFLRGIEKQFLALQKGENEVIPQHLTKTFDEKELIICGLKIDVNMKVNTRLNCT 647
DB 673 WMSKRVGEQFAPIFDGNEELIPQLVAVPDERLELLIGLAVIDCEDMKKHDYNGYQ 732
QY 648 PDSNVKMPKMAVEPFDEERRAKLLQFTGSSRPVLOQKALQGAARLFTTHQIDACT 707
DB 733 ENDQVIAQWFKKINWDSQKARLLQFTGTSRLIPVNGKQSDGDRRPTLRKAGE-A 791
QY 708 NNLPKANTCFNRIDIPYESSYEKLYEKLTAIBETCFPAVE 748
DB 792 NQLPSHTCFNRVLDLPYTDVDSLKQKULTLAVEEVGSGOE 832

```

```

RESULT 12
ID ABJ26104 standard; protein; 869 AA.
XX ABJ26104;
XX
XX 16-APR-2003 (first entry)
XX Aspergillus fumigatus essential gene protein #762.
XX Aspergillus fumigatus essential gene; Aspergillus fumigatus; infection;
XX Fungicide; cytosol; essential; biofilm; antibody; immune response.
XX
XX Aspergillus fumigatus.
XX

```

```

PN WO200286090-A2.
XX
XX 31-OCT-2002.
XX
XX 23-APR-2002; 2002WO-US013142.
XX
XX 23-APR-2001; 2001US-0285637P.
XX
XX 27-APR-2001; 2001US-0287066P.
XX
XX 05-JUN-2001; 2001US-0295890P.
XX
XX 09-JUL-2001; 2001US-0303899P.
XX
XX 31-AUG-2001; 2001US-0316362P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM,
XX WPI; 2003-093124/08.
XX
XX New purified or isolated nucleic acids of essential genes of Aspergillus
XX fumigatus, useful for treating or preventing infections by A. fumigatus,
XX or for treating a non-infectious disease in a subject e.g. cancer.
XX
XX Disclosure; Page; 175p; English.
XX
XX The invention relates to novel purified or isolated nucleic acids of
XX essential genes of Aspergillus fumigatus. The isolated nucleic acids of
XX the invention are used to treat or prevent infections by a pathogenic
XX organism such as A. fumigatus, to treat a non-infectious disease in a
XX subject (e.g. cancer), to prevent or inhibit formation of an object
XX by A. fumigatus, or to prevent or inhibit formation on a surface of a
XX biofilm comprising A. fumigatus. The polynucleotides are useful for
XX expressing recombinant protein for characterization, screening or
XX therapeutic use, as markers for host tissues in which the pathogenic
XX organisms invade or reside, for comparing with the DNA sequence of A.
XX fumigatus to identify duplicated genes or paralogues having the same or
XX similar biochemical activity and/or function, for comparing with DNA
XX sequences of other related or distant pathogenic organisms to identify
XX potential orthologous essential or virulence genes, for selecting and
XX making oligomers for attachment to a nucleic acid array for examination
XX of expression patterns, for raising anti-protein antibodies, as an
XX antigen to raise anti-DNA antibodies or to elicit another immune
XX response, and for identifying polynucleotides encoding the other protein
XX with which binding occurs or to identify inhibitors of the binding
XX interaction. The polypeptides may be used to raise antibodies or to
XX elicit immune response, as a reagent in assays designed to quantitatively
XX determine levels of the protein in biological fluids, as a marker for
XX host tissues in which pathogenic organism invade or reside, and to
XX isolate correlative receptors or ligands in the case of virulence
XX factors. This sequence represents a protein of one of the essential genes
XX of Aspergillus fumigatus of the invention
XX
XX Sequence 869 AA;
XX
XX Query Match 36.0%; Score 1453.5; DB 6; Length 869;
XX Best Local Similarity 36.2%; Pred. No. 6.1e-17;
XX Matches 332; Conservative 114; Mismatches 229; Indels 243; Gaps 21;

```

```

QY 11 PYKRLT-TYLCAKLVKKDFR-----LPPFAKVVVDSGQCH 48
DB 15 PASSRLCVIADGIVKRVFKSYIIAILVRLTSGAFAGPDPFAVYVGGE-QTH 73
QY 49 STDVAKNTLDPKRNQHYDLYIGKSDSVTISVNMNKKIHKKGAGFLGCVRL----- 100
DB 74 TTSVAKKTLNPRYEMFDMRVNEDSIIAIIQFDQKF-KKKDQGLGIVNVRIGVDIDLQ 132
QY 101 -----SNINLK-----DTGYQRDLCTLGNDNDTVRGOIVVSL----- 136
DB 133 MGGDGSLLPIRHSQVSRLLTRLLLEMLTRDLKK--SNDNLVNGKLIINSLTSLST 190
QY 137 -----QSRDRIGTGGVYDCS----- 152
DB 191 PNTNQGALHRSHVQSSSTSGLVLPQVAPSSSHPAAGTAVPDPSANPSINPQVPSSTR 250

```

```

QY 153 -----RLPND---LPGWEERTASGRIOYLNHTRTQ 184
Db 251 PSSTAAPASAGAANSNHSRNTLSFESQGRLPAGWEREDNIGRIYYVDHNTT 310
QY 185 WESPTRPAGSYSPGRLSCFVDENTPIGTINGATGQSSDPLARRVRSQRHNYM-- 242
Db 311 W---TRPSSNYNEHAQ-----RSOREANMQERASHQSRMLP 344
QY 243 -----SRTHLTPP-----DLEGEQSTTQOQVY 268
Db 345 EDRTGANSPLPSSSQAHTPPAGGSAAVSMATGATTAGTGEPLPGWEQSTTEGGRY 404
QY 269 FLHTGTGVSTWHDPRVRDL---SNIN-----CEELGPPPGWEIRNTATGRYV 314
Db 405 FVDHNTRTTWTWDFRQOYIRMYGQVANGNTNTTIOQPVSGQPLPSGWEMLTNTAVY 464
QY 315 FVDHNRRTQFTDPRLSANLHLVLRQNLKQOQOQVSLCPDDECTCTVPRYADVLQ 374
Db 465 FVDHNTKTTWDPRLPSSL-----DQ-----GVFOYKRDPFR 497
QY 375 KXILRQELSQOOP---QAGCRLEVSREIFEEESYQVMKRPKDLKRLMIFRGE 430
Db 498 KLIYFR---SQPALRIMSQCCHKVRNNIFESYAEIMRQASDLKRLMTFQGED 552
QY 431 GLDYGVAREMYLISHENLNPYGLFQYSRDDITTLQINPDSAVNPEHLSYFHFVGRIM 490
Db 553 GLDYGLSGEFPFLSHENENPFYCLFEYSADNTYTLQINPDSGVNPEHLSYFHFVGRIV 612
QY 491 GNAVFHGHYIDGFTLFFYKOLLGKSTLDMELVDPDLHNSLWILENDITGYLDHFC 550
Db 613 GLAIFHRRLDSFFIGAFYKWLKRVSLQMEGVEDLHRLTWTWMDLEGVELTFS 672
QY 551 VEHNAVGEIQLHLEPNKSGIPVNEENKEVYLVNRPFLGIEAOLFALQGFNEVTP 610
Db 673 VDDEKFGERRITDLRGGRDIPVTENKAEYRLTEWKIVKRYVEQGNAMSGFNLIP 732
QY 611 QHLTKTPEKELELITCGKIDVDMKVNRLKCHCTDSIVKMKAVFEPBEKRR 670
Db 733 ADLVAVFERBELITGSLADIYDWMKCHDYKQESDEVIONFWKIVSWDAEKSR 792
QY 671 LLQFVTGSSRVPLQGFKAJQGAAGRLFTIHOIDACTNLEPKATCFNRIDIPYSEYK 730
Db 793 LLQFVTGTSRIPIVNGFKDLQSGDGRRTIEK-SGDPAALPKSHTGNNRDLPPKYET 851
QY 731 LYEXLTAIEETCPAVE 748
Db 852 LEHMKSTAVEETLGFGE 869

RESULT 13
AAM93167
ID AAM93167 standard; protein, 975 AA.
AC AAM93167;
DT 24-MAY-1999 (first entry)
XX
XX Human ZGGBP1 protein.
DE
XX ZGGBP1; bipolar affective disorder type I; human; neurological disorder;
KW regulate; antisense; treatment; disease; stroke; dementia; renal;
XX hypertension; nephrosis; cardiovascular.
XX Homo sapiens.
OS
XX Key location/Qualifiers
FH 1, 975
FT Protein /label= ZGGBP1
FT /note= "Partial sequence"
FT Misc-difference 455
FT /note= "unknown"
XX
XX PN MO906539-AA1.

```

```

XX 11-FEB-1999.
PD
XX 28-JUL-1998; 98WO-GB002259.
PF
XX 01-AUG-1997; 97GB-00016162.
PR
XX (ZENE ) ZENEGA LTD.
PA
XX Flannery AV, Flannegan MCW;
PI
XX WPI; 1999-153769/13.
DR
XX N-PSDB; AAX22554.
XX
XX New isolated gene associated with neurological disorders - used to
XX develop products for treating e.g. stroke, dementia, renal disorders,
XX hypertension or cardiovascular disorders.
XX
XX Claim 1; Page 42-46; 58pp; English.
XX
XX This sequence represents a human ZGGBP1 protein which is associated with
XX affective neurological disorders such as bipolar affective disorder type
XX I. The invention also describes the isolation of the corresponding mouse
XX gene. The ZGGBP1 protein and homologues or fragments may be used to
XX generate substances which selectively bind to it and in so doing regulate
XX the activity of the protein. The antisense DNA or antisense RNA may be of
XX use in the treatment of diseases or disorders in humans in which the over
XX e.g. under-regulated production of the gene product has been implicated,
XX e.g. neurologic (stroke, dementia), renal (hypertension, nephrosis),
XX cardiovascular disorders
XX
XX SQ Sequence 975 AA;
XX
XX Query Match 34.3%; Score 1387; DB 2; Length 975;
XX Best Local Similarity 34.4%; Pred. No. 1,4e-120;
XX Matches 330; Conservative 125; Mismatches 247; Indels 258; Gaps 26;
XX
XX 14 LRLTVLCANLVKXDFRLLDPFAVY---VVDGSGCH--SDTYKNTLDPKMNQHYLY 68
XX 42 LRVKVVSGIDLAKKQIFGASDPYVLSLVADENELAVQTKTKILNPKANNEEFYFR 101
XX 69 IGKSD-SVTISVWNEKTKHKQAGFLCCVR-----LLSNAIN 105
XX 102 VNPSNRHLFEVFDENRLTRD---FLQGVVPLSLHPTEDPTMERPYTFKDFLLRPSH 158
XX 106 RLKDTGYORLDCXKLPNDNDTVRGQIVSVLSQSRDRIQTGGQVDCSLRFPND----- 158
XX 159 KSRVKGFLRLKMAVYPRKG-----QDEENSDQDDMEHGWVVD-----SNDASQHOE 208
XX 159 -----LPGWEERTASGRIOYLNHTRTQWERPT----- 189
XX 209 ELPPPLPPGWEKVDNIGRTYVYVNNRRTQWRHPSLMDVSSSDNNIRQINQEAHR 268
XX 190 -----RPASEYSP-----GRPLSCFVDENTPISGIN-----GATGQSSDPR-- 227
XX 269 FRSRHHISEDLPEPSEGDVPEPWEET-ISEEVNIAGSLGVLPAPPASGSRISPOEL 327
XX 228 -----LAER-----RVRS-----QHRRYMYSRTHLHTPPDL 253
XX 328 SEELSRRLQITPDSNGEQFSSLIRBPSRSRSGSVMDAVAEOHLPPPSVAYHTHTPGL 387
XX 254 PEGYQRTQOQGVYFLHTQGVSTWHP-----RVPRDL 288
XX 388 PSGEERKDAKGRITYVNNHNRRTTWTWTRPIQLAEDGASGATNSNNHLTPQIRRRSL 447
XX 289 SNINCEELGP-----
XX 448 SSPVYTLXAPLEGAKDSPVARAVDTLSNPQSPQPSYNSPKQHKYTKQSLPFGWEMRI 507
XX 308 TATGRVYFVDHNRNTTQFTDPRLSANLHLVLRQNLKQOQOQVSLCPDD----- 359
XX 508 APNGRPFIDHNTKTTWEDPRLKFPVHM-----RSKXSLNENDLGFLPPGW 554
Db

```

```

QY 360 -----TECLTVP--YKRDIVOKLILQELSQOQPC 389
DB 555 EERHIDGRTFYIDHNSKITQWEDPLQNPATIGPVPSPRFKQYDFRKKLKKPADI 614
QY 390 AGCHREIVSREEEFEESSYQVMKRPDKMK-RLMIKFRGEGLDYGVARFWLYLSHE 448
DB 615 PNFPEWKLHNNIFEESSYRIRMSVKRPDLKARLWIEFESEKGLDYGVARFWFLSKXE 674
QY 449 MNPVYGLQYASDDIYTLQINPDASV-NPEHLSYFHFVGRIMGVAHFHGHYIDGGFTLP 507
DB 675 MNPVYGLQYASDDIYTLQINPDASV-NPEHLSYFHFVGRIMGVAHFHGHYIDGGFTLP 734
QY 508 FYKOLIGKSTLDDMLVDPDLANSJWLLENDITGVLDHTPCVENHAYGEIIQHELKN 567
DB 735 FYKWLIGKSTLDDMLVDPDLANSJWLLENDITGVLDHTPCVENHAYGEIIQHELKN 793
QY 568 GSKIPVNEBKKEVYVLYNNRFLRGIEQFLALQKGFNEVPOHLLKTFDEKELEIIC 627
DB 794 GSEIWTNENKREYIDLVIQWRFVNEVQKMAFLGFTLELLPIDIKLFDENELLELMC 853
QY 628 GCGKIDVNDKVTNRLK--HCTPDSNIVKFWKAVEFPDEERRARLLQFVTGSSRYPLQG 685
DB 854 GLGDVDVNDKROHSIKNGYC-PNHVVIQWFWKAVLLMDAKRIRLLQFVTGSRYPNMG 912
QY 686 FKALQGAAGRLFTTHQIDACTNNLPKATCFNRIDIPPEYSEYKLYEKLTLAIEETGCF 745
DB 913 FAEYLSNGPQLFTIEQWGS-PEKLPRAHCFNRDLPPYEYFEDIREKLLMAVENAQA 971

RESULT 14
ADB73455 standard; protein, 911 AA.
XX ADB73455;
XX AC ADB73455;
XX 04-DEC-2003 (first entry)
XX DE Prostate cancer marker protein.
XX DE Prostate cancer marker protein.
XX KM Prostate; cancer; cytostatic; gene therapy; marker.
XX OS Homo sapiens.
XX PN W02003009814-A2.
XX PD 06-FEB-2003.
XX PF 25-JUL-2002; 2002MO-US023913.
XX PR 25-JUL-2001; 2001US-0307982P.
XX PR 22-AUG-2001; 2001US-0314356P.
XX PR 23-SEP-2001; 2001US-0325020P.
XX PR 12-DEC-2001; 2001US-0341746P.
XX PR 05-MAR-2002; 2002US-0362158P.
XX PA (MILL-) MILLENNIUM PHARM INC.
XX PI Schlegel R, Monahan JE, Endege WO, Gamavapapu M, Gorbacheva B,
XX PI Heerth S, Kamakar S, Womsey AM, Glatc K, Zhao X, Anderson D,
XX DR WPI; 2003-248033/24.
XX PT New nucleic acid molecule, useful for diagnosing or treating prostate
XX cancer.
XX PS Claim 4; SEQ ID NO 279; 999P; English.
XX CC The invention relates to newly discovered cancer markers associated with
XX CC the cancerous state of prostate cells. Also disclosed is a method of
XX CC assessing whether a patient is afflicted with prostate cancer. The method
XX CC of the invention involves assessing whether a patient is afflicted with
XX CC prostate cancer by comparing the level of expression of a marker in a
XX CC patient sample and the normal level of expression of the marker in a

```

```

CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB75177-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 911 AA;
XX
QY Query Match 33.9%; Score 1370.5; DB 7; Length 911;
QY Best Local Similarity 36.7%; Pred. No. 4.3e-119;
QY Matches 324; Conservative 119; Mismatches 234; Indels 207; Gaps 28;
DB 24 LYKSDFFELP--GPPAKV--VVDGSGGCH--STDVKNLDPKRNQHYDLYGSD-S 74
DB 69 LPTSDELFLPGLCPYKLSLYVADENBELALVOTKIKTLNKNMEEFYFRVPSNHR 128
QY 75 VTISVNNHKKLHKKGAGFLGCVR-----LISNAINRLKDTGY 112
DB 129 LLEFVFDENRLTRDD--FLGQVDVPLSHLPTEDPTMERPYTFKDFLLRPSHRKRVKGF 185
QY 113 QRLDCLKGPNDNTVAGQIVVSLQSRDRIGTGQVYDCSLFEND-----L 159
DB 186 LRLKAVYPPKKG--GDEENSDQRDMEHGEVVD--SNDASAGQHELPPEPL 235
QY 160 PDGWEERRTASGRIOYLNHITRTTOWERPT-----RPA 192
DB 236 PPGMEKXYDNGRTYVYVNNHNRITQWHRPSLMDVSSSDNNRIQINQEAHRRFRSRHI 295
QY 193 SEYSSP-----GRPLSCFVDENTPIISGTN-----GATGQSSDPR-LAE--RR 232
DB 296 SEDLEPSEEGDYPEPWT-ISEEVNIAGDSLALPPPASGSRISPOELSELSRR 354
QY 233 VR-----SORHNNYSRTHL-----HTPPD----- 252
DB 355 LQITPDSNGEQFSSLIQRPSSRLSCSVTDAVAEQHLPFGAKDSPVRAVKDTLSNPQ 414
QY 253 -----LPEGYEQRTQGGVYFHTQTVSYTMHDPVP-----RD 287
DB 415 SPQEPSPVSPKQHKVQTSFLPPGMEWRIAENGPFFLDNHTKTYTIEDPRLKPPVHRS 474
QY 268 LSNICEELGRLPPGMEIRNTATGVRVVDHNNRTQFTDPRLSANLHLVLRNQNLQDQ 347
DB 475 KTSLNPNDLGPPGMEIRHLDGRTFYIDHNSKITQWEDRL-----QN----- 519
QY 348 QQQQVVSCLPDDTECLTVP--YKRDIVOKLILQELSQOQPAAGHRIEVSREEEFEE 405
DB 520 -----PATGPAPVPSREPFQKYDYFRKKLKKPADIPNPEMLHNNIPEE 566
QY 406 SYRQVMKRPDKMK-RLMIKFRGEGLDYGVARFWLYLSHMLNAYYGLFOYSRDI 464
DB 567 SYRIRMSVYRPDVLKARLWIEFESEKGLDYGVARFWFLSKMEFNYVYGLFESADQN 626
QY 465 YTLQINPDASV-NPEHLSYFHFVGRIMGVAHFHGHYIDGGFTLPFYKOLIGKSTLDDME 523
DB 627 YTLQINPDASV-NPEHLSYFHFVGRIMGVAHFHGHYIDGGFTLPFYKOLIGKSTLDDME 686
QY 524 LYDPLANSJWLLENDITGVLDHTPCVENHAYGEIIQHELKNKSKSPVNEBKKEVYR 583-
DB 687 SVDSSEYNSLKLWILENDITGVLDHTPCVENHAYGEIIQHELKNKSKSPVNEBKKEVYR 745
QY 584 LYVNNRFLRGIEQFLALQKGFNEVPOHLLKTFDEKELEIICGLKIDVNDKVTNRL 643
DB 746 LYVNNRFLRGIEQFLALQKGFNEVPOHLLKTFDEKELEIICGLKIDVNDKVTNRL 805
QY 644 K--HCTPDSNIVKFWKAVEFPDEERRARLLQFVTGSSRYPLQGFKALQGAAGRLFTTH 701
DB 806 KNGYC-PNHVVIQWFWKAVLLMDAKRIRLLQFVTGSRYPNMGFAEYLSNGPQLFTIE 864
QY 702 QIDACTNNLPKATCFNRIDIPPEYSEYKLYEKLTLAIEETGCF 745

```

DB 865 QWGS-PEKLPRAHTCFNRLLDPYETFEEDLQEKLMAVEANAQSF 907

RESULT 15

ID AAW36797 standard; peptide; 724 AA.

AC AAW36797;

XX 23-APR-1998 (first entry)

DE Novel human gene, designated WMP4.

XX Peptide recognition unit; YAP WW domain binding protein; WBP-1; WBP-2;

KW WW domain; cell signalling; growth regulation; cytoskeleton organisation;

KW targeted drug screening; modulator; WW domain interaction; WMP4.

XX Homo sapiens.

XX Key Location/Qualifiers

FT misc\_difference 1..3

FT /note= "the nucleotides encoding these amino acids are

FT not given in the specification"

FT Domain

FT 140..163

FT /note= "Claimed (claim 49) WW Domain 1"

FT 252..277

FT /note= "Claimed (claim 49) WW domain 3"

FT 303..328

FT /note= "Claimed (claim 49) WW domain 3"

FT 618..724

FT /note= "Claimed (claim 90) HECT domain"

XX MO9737223-A1.

XX 09-OCT-1997.

XX 03-APR-1997; 97WO-US005547.

XX 03-APR-1996; 96US-00630916.

XX {CYTO-} CYTOGEN CORP.

XX {UNYC-} UNIV NORTH CAROLINA.

XX Pirozzi G, Kay BK, Fowlkes DM;

XX WPI: 1997-503234/46.

XX N-PSDB; AAT95700.

XX Identifying cell signalling and growth regulatory polypeptides by

PT reaction with multivalent recognition complex - polypeptides are useful

PT in targeted drug selection.

XX Claim 48; Fig 23; 220p; English.

XX The present sequence represents a novel protein WMP4. The WMP4 gene was

CC identified and isolated from a cDNA expression library generated from

CC LNCaP prostate cancer cell line, using peptides AAW3603-64. These

CC peptide recognition units are based on the sequences of WW domain binding

CC domains of the alpha and gamma subunits of epithelial sodium channel

CC protein. The WW domain is a small functional domain found in a large

CC number of proteins from a variety of species including humans, nematodes

CC and yeast. Its name is derived from the observation that two tryptophan

CC residues, one in the amino terminal portion of the WW domain and one in

CC the carboxyl terminal portion, are conserved. Most proteins containing WW

CC domains have a function involving cell signalling and growth regulation

CC or the organisation of the cytoskeleton. Polypeptides containing a WW

CC domain are identified by treating a multivalent recognition unit complex

CC that has selective binding affinity for a WW domain, with many

CC polypeptides and identifying those with selective affinity for the

CC complex. Proteins containing WW domains are used for targeted drug

CC screening, i.e. to identify potential modulators of specific WW domain

CC interactions. The valency of the recognition unit is important in

CC determining specificity of interaction with WW domains. In multivalent

CC form specificity is relaxed, but not lost, so proteins containing WW

CC domains similar, but not identical, to the sequence of the peptides

CC target WW can be detected, including new polypeptides

XX

SQ Sequence 724 AA;

Query Match 33.6%; Score 1355.5; DB 2; Length 724;

Best Local Similarity 43.7%; Pred. No. 7.7e-118;

Matches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;

QY 137 QSRDRIGTGQGVDCSRLEFDN---DLDPGEERRTASGRIOYLNHITTTQWBERTPAS 193

DB 111 RARSSTVATGGEPTPSVAVHTTPTGLPGMEERRKDAKRTYVNNHNTTWTTPIMQIA 170

QY 194 ESSSG-----RPLSC---FVDENTPISGTMGACGSSDRLEARRVS 235

DB 171 EDGASGASNNNNHILIEPQIRPRSLSPFTVLSAPAGADSPVRRAVKDTLSNPSQPQ 230

QY 236 QHRNRYMSRTHLHPEDLPEGYEQTTOGQGVYFLHTQTGVSTWHPRPV-----RDLN 290

DB 231 PEPYNSPKPQHKVTSFPLPGMEKRIAPNGSPFFIDHTKTTWEDPLKFPVHMRKTS 290

QY 291 INCEELGFLPPGMEIRNTATGRVYFVDHNNNTTQFTPPRLSANHLVLRNOLKDOQQQ 350

DB 291 INPNDLGPLPGEERRIHLDRFTFYIDNSKITQWEDPRL-----QN----- 332

QY 351 QVSLCPDDECTATVPR--YKRDVQKLIIRQELSCQQQPOAGCRIEVSREIFESYSR 408

DB 333 -----PAITPAVPYGRPEKQKDYRKKLTKKPADIPIINFEMKLNHNHNFESYSR 382

QY 409 QVMKRPDLWK-RIMKFRGEESGIDYGVARENLVYLISHMLNPYGLFOYSRDIYTL 467

DB 383 RIMSVKRDVVKARLWIEFSEKGLDYGVARERWFFLSKEMFNPHYGLFEYSATDNYTL 442

QY 468 QINPDAV-NPEHLSYFHFVGRIMGMAVFHGHIYDGGFTLPFYKQLIGKSTLTDMLVD 526

DB 443 QINPMSGCNEPDHLSYFTFGKVAGLAVFHKLLDGFIRPFYKMLGKQITLNDMSVD 502

QY 527 PDLNSLVILLENDITGLDHTFCVENNAVGEIIOHEIKPNGKSIPIVENEKKYVRLVY 586

DB 503 SEYNSLKWILENDPT-ELDLMFCIDEENFQYQVQLKPGSEIMVTNEKREYIDLVI 561

QY 587 NWRPLRGLEAQLQKGFNEVIFQHLKTPDELELEITCGIKIVNPKNTRLK-- 644

DB 562 QRFVNRVQKQNNAPLESGFTLPLDLKTFDENLELMLKGLDGDVNDMRQHSIYKNG 621

QY 645 HCTPDSNIVKMFMAVEFFDEERRARLLQFVTGSSRPVLOGFKALQGAAGRLLFTIHOID 704

DB 622 YC-PHNPVIOQFWKAVLMDAEKRIILLQFVTGSRVPMNGFARLYGNSGQLFTISQWG 680

QY 705 ACTNNLPRAHTCFNRLLDPYETFEEDLQEKLMAVEANAQSF 745

DB 681 S-PEKLPRAHTCFNRLLDPYETFEEDLQEKLMAVEANAQSF 720

Search completed: September 21, 2004, 07:43:44

Job time : 98.5466 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: September 21, 2004, 07:28:49 ; Search time 89.4534 Seconds  
(with hour alignments)  
2283.667 Million cell updates/sec

Title: US-10-009-945-2

Percent score: 3884  
Sequence: 1 GGSISIKIRLTVCARLAKK.....EKLYEKLTAVENTGFAVE 723

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: A\_Geneseq\_29Jan04:\*

1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001s:.\*  
5: Geneseq2002s:.\*  
6: Geneseq2003as:.\*  
7: Geneseq2003bs:.\*  
8: Geneseq2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3884	100.0	723	AA031476	Aab31476 Amino aci
2	3878	99.8	722	AA032724	Aae32724 E3 ubiqui
3	3869	99.6	722	AA079861	Aam79861 Human pro
4	3001	77.3	748	AA013384	Aaw13384 Human pro
5	3001	77.3	748	AA013384	Aab31477 Amino aci
6	2996	77.1	804	AAU19610	Aau19610 Human dia
7	2236	57.6	514	AAU87301	Aau87301 Novel cen
8	2069	53.3	1035	AA061120	Aab61120 Drosophi
9	1614.5	41.6	335	AA078877	Aam78877 Human pro
10	1461	37.6	766	AA013385	Aaw13385 Human pro
11	1436.5	37.0	832	ABP73459	Abp73459 Candida a
12	1345	34.6	869	ABU26104	Abj26104 Aspergill
13	1318	33.9	854	AA030949	Aay30949 Murine E3
14	1301	33.5	852	AA030948	Aay30948 Human E3
15	1288.5	33.2	903	AB064210	Ab064210 Angiogene
16	1259	32.4	927	AA025170	Aay25170 Human NTA
17	1259	32.4	927	AA032725	Aae32725 Human NED
18	1249.5	32.2	911	ADB75455	Adb75455 Prostate
19	1245.5	32.1	739	ABG16477	Abg16477 Novel hum
20	1245.5	32.1	739	ABG68173	Abg68173 Atrophin-
21	1245.5	32.1	739	AA032720	Aae32720 Atrophin-
22	1245.5	32.1	739	ABR41097	Ab41097 Human atr
23	1245.5	32.1	739	ADB98717	Ad98717 Human pro
24	1245.5	32.1	739	ADB82513	Ad82513 Human pro
25	1245.5	32.1	752	ADB77070	Ad77070 Human pro

26	1245	32.1	975	2	AA093167
27	1239	31.9	922	4	AA050049
28	1239	31.9	522	5	AB050708
29	1239	31.9	949	4	AB059631
30	1239	31.9	949	4	AB059836
31	1228.5	31.6	995	6	AA032719
32	1228.5	31.6	995	6	AA032719
33	1228.5	31.6	995	6	AA032719
34	1228.5	31.6	995	6	AA032719
35	1228.5	31.6	995	6	AA032719
36	1227	31.6	725	2	AA036797
37	1227	31.6	725	2	AA036797
38	1226.5	31.6	759	6	AB05504
39	1226	31.6	854	7	ADB75453
40	1225	31.5	870	4	AA050495
41	1225	31.5	870	6	AA032722
42	1225	31.5	906	7	ADB49242
43	1223	31.5	854	5	AAU77715
44	1223	31.5	854	6	ABP58332
45	1212	31.2	474	2	AAU10943

## ALIGNMENTS

RESULT 1

ID AAB31476 standard; protein; 723 AA.

AC AAB31476;

DT 20-APR-2001 (first entry)

DE Amino acid sequence of a human Smurf1 polypeptide.

KW Smurf1; Smurf2; Smad signal transduction; bone morphogenic protein; BMP; transforming growth factor-beta; human; TGF-beta; Chondrogenesis; osteogenesis; blood differentiation; cartilage formation; hair growth; neural tube patterning; retinal development; heart induction; morphogenesis; tooth formation; gamete formation.

KM Homo sapiens.

OS Homo sapiens.

XX WO200077168-A2.

XX 12-JUN-2000; 2000WO-US016250.

XX 11-JUN-1999; 99US-0138969P.

XX (UNYNY) UNIV NEW YORK STATE RES FOUND.  
(HSCR-) HSC RES & DEV LP.

XX Thomsen GH, Wraana J;

XX WPT. 2001-071267/08.

XX N-PDB; AAF24852.

PT Novel isolated Smurf protein useful for inhibiting bone morphogenic protein or tumor growth factor-beta activation pathway, for treating cancer and to block osteogenesis, hair growth, tooth formation.

PS Claim 6; Fig 10; 107pp; English.

CC The present sequence represents a human Smurf1 polypeptide. The specification also describes a Smurf2 polypeptide. Smurf polypeptides are negative regulators of Smad signal transduction, and antagonists of bone morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta) signaling pathway. Expression of Smurf1 in a cell is useful for inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf polypeptides are useful for blocking chondrogenesis, osteogenesis, blood differentiation, cartilage formation, neural tube patterning, retinal

CC development, heart induction and morphogenesis, hair growth, tooth  
 CC formation, gamete formation and a wide variety of tissue and organ  
 CC formation processes, and hinder the regeneration, growth, maintenance,  
 CC etc., of bone and other tissues that are dependent on the BMP pathway.  
 CC The polypeptide is useful for screening for various drugs and/or  
 CC antibodies that can either enhance the BMP pathway, or inhibit it  
 CC

SQ Sequence 723 AA;

Query Match 100.0%; Score 3884; DB 4; Length 723;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSSSIKIRLTVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTDTVKTLDPKXNQHLYD 60  
 DB 1 GSSSIKIRLTVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTDTVKTLDPKXNQHLYD 60  
 QY 61 YVGKTDSTITISVWNHKKIHKKGAGFLGCVLLSNALSRLKDTGYQRDLCKLNPSTDA 120  
 DB 61 YVGKTDSTITISVWNHKKIHKKGAGFLGCVLLSNALSRLKDTGYQRDLCKLNPSTDA 120  
 QY 121 VVGQIVLSLQTRDRIGTGGVYVDCRGLLENNGTYEEDSGPRPLSCMEBPAPYTDSTGA 180  
 DB 121 VVGQIVLSLQTRDRIGTGGVYVDCRGLLENNGTYEEDSGPRPLSCMEBPAPYTDSTGA 180  
 QY 181 AAGGNCRFVESPQDQRLQARLNPDPVRSLSLQTPQNRPHGQSPPLPSGEQRTTVQG 240  
 DB 181 AAGGNCRFVESPQDQRLQARLNPDPVRSLSLQTPQNRPHGQSPPLPSGEQRTTVQG 240  
 QY 241 QVFLHTQGVSTWHDPRIPRLDLSVNCDELGPPLPGMEVRSIVSGRIYVDHNNRTQFT 300  
 DB 241 QVFLHTQGVSTWHDPRIPRLDLSVNCDELGPPLPGMEVRSIVSGRIYVDHNNRTQFT 300  
 QY 301 TDPRLHIMNHQCOLKEPSCPLPLPSESGLEDEELPAQRYERDLVQKLKLRHELSIQP 360  
 DB 301 TDPRLHIMNHQCOLKEPSCPLPLPSESGLEDEELPAQRYERDLVQKLKLRHELSIQP 360  
 QY 361 QAGHCRLEVSREELFEESYRQIMKRRPDLKRLMKVFGEGLDYGVAREWLYLICHE 420  
 DB 361 QAGHCRLEVSREELFEESYRQIMKRRPDLKRLMKVFGEGLDYGVAREWLYLICHE 420  
 QY 421 MLNPPYGLFQYSTDNIVMLQINPDSSINPDLSYFHFVGRIMGLAVFHGHYNGFTVP 480  
 DB 421 MLNPPYGLFQYSTDNIVMLQINPDSSINPDLSYFHFVGRIMGLAVFHGHYNGFTVP 480  
 QY 481 YKQLISPRQLSDLESYDPELHNSLVWILENDITPVLDTFCVEHNAFGRIILQHELPNG 540  
 DB 481 YKQLISPRQLSDLESYDPELHNSLVWILENDITPVLDTFCVEHNAFGRIILQHELPNG 540  
 QY 541 RNVPTEENKKEVYRLVYVNRPMRGIEAOFALQKGFNELIPQHLKPEQKLELITGG 600  
 DB 541 RNVPTEENKKEVYRLVYVNRPMRGIEAOFALQKGFNELIPQHLKPEQKLELITGG 600  
 QY 601 LDKTLDLNDKSNLRLKHCVAADSNIVRFWQAVETFEDEERARLLQFVTSRVLQGFKA 660  
 DB 601 LDKTLDLNDKSNLRLKHCVAADSNIVRFWQAVETFEDEERARLLQFVTSRVLQGFKA 660  
 QY 661 LQGSTGAGRLFTFIHIDANTDNLPRATCCNRIDIPPESEYKLYEKLITAVEETGCF 720  
 DB 661 LQGSTGAGRLFTFIHIDANTDNLPRATCCNRIDIPPESEYKLYEKLITAVEETGCF 720  
 QY 721 AVE 723  
 DB 721 AVE 723  
 QY 721 AVE 723  
 DB 721 AVE 723

RESULT 2

AAE32724 ID AAE32724 standard; protein; 722 AA.

XX AAE32724;

DT 24-MAR-2003 (first entry)

XX E3 ubiquitin ligase SMURP1 protein.  
 DE  
 XX Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;  
 KM polioyellitis; HIV; measles; protein therapy; E3 ubiquitin ligase;  
 XX enzyme.  
 KM

OS Unidentified.

XX WO200290549-A2.

XX 14-NOV-2002.

XX 12-MAR-2002; 2002WO-1B002106.

XX 12-MAR-2001; 2001US-0275224P.

XX 31-JUL-2001; 2001US-0308958P.

XX 07-DEC-2001; 2001US-0340170P.

XX (PROT-) PROTEOLOGICS LTD.

XX Greener T, Moskowitz H, Reiss Y, Alroy I;

XX MPI; 2003-111976/10.

XX N-PSDB; AAD50458.

XX New protein complex comprising HECT-RCCL, viral maturation scaffolding  
 PT protein (VWSP), and/or HIV gag protein, useful for treating viral  
 PT infections, such as lymphosarcoma, HIV, hepatitis, polioyellitis, measles,  
 PT or Ebola.

XX Disclosure; Fig 14; 150pp; English.

XX The invention relates to a method for modulation of viral maturation. The  
 CC invention also provides an isolated protein complex comprising a HECT-  
 CC RCCL selected from HECT-WM, HECT-RCCL, Gag protein, Gag late domain, P13,  
 CC actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIV,  
 CC GTPase, E2 enzyme, Esg101, cullin, HECT1, HECT2, HECT3, Nedd4-like  
 CC protein or clathrin. The complexes, proteins, antibodies and methods are  
 CC useful for treating viral infections, such as lymphosarcoma, human  
 CC immunodeficiency virus (HIV), hepatitis, polioyellitis, measles, or Ebola  
 CC and for inhibiting budding in a subject. They are also useful in  
 CC diagnostic assays for determining whether a cell is infected with a virus  
 CC and for characterizing the nature, progression and/or infectivity of the  
 CC infection. The invention is also useful in protein therapy. The present  
 CC sequence is E3 ubiquitin ligase SMURP1 protein used to illustrate the  
 CC method of the invention

SQ Sequence 722 AA;

Query Match 99.8%; Score 3878; DB 6; Length 722;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSSSIKIRLTVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTDTVKTLDPKXNQHLYD 61  
 DB 1 GSSSIKIRLTVLCANLAKKDFRLLPDPFAKIVVDSGQCHSTDTVKTLDPKXNQHLYD 60  
 QY 62 YVGKTDSTITISVWNHKKIHKKGAGFLGCVLLSNALSRLKDTGYQRDLCKLNPSTDAV 121  
 DB 62 YVGKTDSTITISVWNHKKIHKKGAGFLGCVLLSNALSRLKDTGYQRDLCKLNPSTDAV 120  
 QY 122 VVGQIVLSLQTRDRIGTGGVYVDCRGLLENNGTYEEDSGPRPLSCMEBPAPYTDSTGAA 181  
 DB 122 VVGQIVLSLQTRDRIGTGGVYVDCRGLLENNGTYEEDSGPRPLSCMEBPAPYTDSTGAA 180  
 QY 181 AAGGNCRFVESPQDQRLQARLNPDPVRSLSLQTPQNRPHGQSPPLPSGEQRTTVQG 241  
 DB 181 AAGGNCRFVESPQDQRLQARLNPDPVRSLSLQTPQNRPHGQSPPLPSGEQRTTVQG 240  
 QY 242 VYFLHTQGVSTWHDPRIPRLDLSVNCDELGPPLPGMEVRSIVSGRIYVDHNNRTQFT 301  
 DB 242 VYFLHTQGVSTWHDPRIPRLDLSVNCDELGPPLPGMEVRSIVSGRIYVDHNNRTQFT 300

QY 302 DPLHHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLYOKLKYLRHLSLQOPQ 361  
 DB 301 DPLHHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLYOKLKYLRHLSLQOPQ 360  
 QY 362 AGHCRLEVSREELFEESYRQIMMRPKDLKKRLMVKFGEGLDYGVARERWLYLLCHEM 421  
 DB 361 AGHCRLEVSREELFEESYRQIMMRPKDLKKRLMVKFGEGLDYGVARERWLYLLCHEM 420  
 QY 422 LNPYYGLFQYSTDNITMQLINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGFTVPY 481  
 DB 421 LNPYYGLFQYSTDNITMQLINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGFTVPY 480  
 QY 482 KOLLGKPIQLSDLESVDPELHKSILWILENDITPVLDFHFCVEHNAFGRILLOHELKPNR 541  
 DB 481 KOLLGKPIQLSDLESVDPELHKSILWILENDITPVLDFHFCVEHNAFGRILLOHELKPNR 540  
 QY 542 NVPVTEENKKEYVRLVYVMRFMRGIEAQLALOKGFNELIPQHLKPFQKLELIIIGL 601  
 DB 541 NVPVTEENKKEYVRLVYVMRFMRGIEAQLALOKGFNELIPQHLKPFQKLELIIIGL 600  
 QY 602 DKIDLNDKSNTRLKQCVADSNIVRMPQAVTEFDEERRARLLQVYTGSTRVPLQGFAL 661  
 DB 601 DKIDLNDKSNTRLKQCVADSNIVRMPQAVTEFDEERRARLLQVYTGSTRVPLQGFAL 660  
 QY 662 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPPESEYKLYEKLITAVEETCGFA 721  
 DB 661 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPPESEYKLYEKLITAVEETCGFA 720  
 QY 722 VE 723  
 DB 721 VE 722

RESULT 3  
 AAM79861  
 ID AAM79861 standard; protein: 722 AA.  
 XX  
 AC AAM79861;  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human protein SEQ ID NO 3507.  
 XX  
 KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KM nervous system disorder; arthritis; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 FN W0200157190-A2.  
 PD 09-AUG-2001.  
 XX  
 PE 05-FEB-2001; 2001WO-US04098.  
 XX  
 PR 03-FEB-2000; 2000US-00496914.  
 PR 27-APR-2000; 2000US-00560875.  
 PR 20-JUN-2000; 2000US-00598075.  
 PR 19-JUL-2000; 2000US-00620325.  
 PR 01-SEP-2000; 2000US-00654936.  
 PR 15-SEP-2000; 2000US-00663561.  
 PR 20-OCT-2000; 2000US-00693325.  
 PR 30-NOV-2000; 2000US-00728422.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 PI Tang YT, Liu C, Dzmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
 PI Ma Y, Zhao QA, Wang J, Zhang J, Ren F, Chen R, Wang ZW,  
 PI Xue AJ, Yang Y, Wejhtman T, Goodrich R;  
 XX WPI; 2001-476283/51.

DR N-PSDB; AAK52994.  
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 PT in diagnosis and gene therapy.  
 PT  
 XX  
 PS Claim 20; Page 362-363; 6221pp; English.  
 XX  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, immunomodulatory activity and  
 CC activity/inhibit activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication  
 CC  
 SQ Sequence 722 AA;  
 Query Match 99.6%; Score 3869; DB 4; Length 722;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 GSSIKIRLTVLCAKNLAKKDFRLLDPFAKIVVDGSGGCHSTDTYKNTLDRKNOHYDLY 61  
 DB 1 GSSIKIRLTVLCAKNLAKKDFRLLDPFAKIVVDGSGGCHSTDTYKNTLDRKNOHYDLY 60  
 QY 62 VKTDSITISVWNHKKIKKKGAGTGCVRLLSNASISLXDTGVRDLCLNPSDTPAY 121  
 DB 61 VKTDSITISVWNHKKIKKKGAGTGCVRLLSNASISLXDTGVRDLCLNPSDTPAY 120  
 QY 122 RQIIVVSLQTRIDRIGTGSVVDGRLNENGTVEYDSGGRPLSGFMEBPAPYTDSTGAA 181  
 DB 121 RQIIVVSLQTRIDRIGTGSVVDGRLNENGTVEYDSGGRPLSGFMEBPAPYTDSTGAA 180  
 QY 182 AGGNCRFVESPSQORLQAOQLRNPVDRGSLQTPQNPFGHQSLEBGEYQRTVVGQ 241  
 DB 181 AGGNCRFVESPSQORLQAOQLRNPVDRGSLQTPQNPFGHQSLEBGEYQRTVVGQ 240  
 QY 242 VYFLHTQGVSTWHPRLPRDLNSVNCDELGLPFGMEVRSVSGRIYFVDHNRRTQFT 301  
 DB 241 VYFLHTQGVSTWHPRLPRDLNSVNCDELGLPFGMEVRSVSGRIYFVDHNRRTQFT 300  
 QY 302 DPLHHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLYOKLKYLRHLSLQOPQ 361  
 DB 301 DPLHHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQRYERDLYOKLKYLRHLSLQOPQ 360  
 QY 362 AGHCRLEVSREELFEESYRQIMMRPKDLKKRLMVKFGEGLDYGVARERWLYLLCHEM 421  
 DB 361 AGHCRLEVSREELFEESYRQIMMRPKDLKKRLMVKFGEGLDYGVARERWLYLLCHEM 420  
 QY 422 LNPYYGLFQYSTDNITMQLINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGFTVPY 481  
 DB 421 LNPYYGLFQYSTDNITMQLINPDSSINPHLSYFHFVGRIMGLAVFHGYINGGFTVPY 480  
 QY 482 KOLLGKPIQLSDLESVDPELHKSILWILENDITPVLDFHFCVEHNAFGRILLOHELKPNR 541  
 DB 481 KOLLGKPIQLSDLESVDPELHKSILWILENDITPVLDFHFCVEHNAFGRILLOHELKPNR 540  
 QY 542 NVPVTEENKKEYVRLVYVMRFMRGIEAQLALOKGFNELIPQHLKPFQKLELIIIGL 601  
 DB 541 NVPVTEENKKEYVRLVYVMRFMRGIEAQLALOKGFNELIPQHLKPFQKLELIIIGL 600  
 QY 602 DKIDLNDKSNTRLKQCVADSNIVRMPQAVTEFDEERRARLLQVYTGSTRVPLQGFAL 661  
 DB 601 DKIDLNDKSNTRLKQCVADSNIVRMPQAVTEFDEERRARLLQVYTGSTRVPLQGFAL 660  
 QY 662 QGSTGAAGPRLFTIHLIDANTDNLPRKATCFNRIDIPPESEYKLYEKLITAVEETCGFA 721

Db 661 QGSGAAGRLFTLHLIDANTIDNLKAKTCENR.LDIPPESEYKYLTAVEETCGFA 720  
 QY 722 VE 723  
 Db 721 VE 722

RESULT 4  
 AAM13384  
 ID AAM13384 standard; protein; 748 AA.  
 AC AAM13384;

DT 10-JUL-1997 (first entry)

DE Human protein ubiquitin ligase pub1.

KM Protein ubiquitin ligase; pub1; cdc25 phosphatase; CDK kinase; p53;  
 cell cycle; transgenic animal.

OS Homo sapiens.

PN WO9712962-A1.

PD 10-APR-1997.

PF 04-OCT-1996; 96MO-US015930.

PR 04-OCT-1996; 95US-00539205.

PA (COLD-) COLD SPRING HARBOR LAB.

PI Beach D, Caligiuri M, Nefsky B;

DR WPI: 1997-226206/20.

DR N-PSDB; AAM17040.

PT Human and Saccharomyces pombe protein ubiquitin ligase(s) - involved in  
 PT cdc25 phosphatase and p53 ubiquitination, and regulate cell growth and  
 PT proliferation.

PS Claim 1; Page 74-77; 108pp; English.

CC Human protein ubiquitin ligases pub1 (AAM13384), pub2 (AAM13385) and pub3  
 CC (AAM13386) are homologues of fission yeast pub1 (AAM13387) and were  
 CC identified from cDNA clones (AAT47040-42) obt'd. e.g. from a keratinocyte  
 CC cDNA library. Pub polypeptides can be produced in transfected host cells.  
 CC They can control the steady state level of cdc25 phosphatase, the degree  
 CC of CDK kinase, (e.g. cdc2) dephosphorylation and the steady state level of  
 CC p53 (controlling the degree of cell cycle regulation of p53). They can  
 CC regulate cell or tissue differentiation, or cell growth or proliferation  
 CC by affecting other proteins, can be a specific (ant)agonist of wild-type  
 CC protein function and may be used as immunogens to elicit a specific  
 CC immune residue

SO Sequence 748 AA;

Query Match 77.3%; Score 3001; DB 2; Length 748;  
 Best Local Similarity 74.6%; Pred. No. 1.5e-285;  
 Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

QY 5 IKRLTYCACNLAKKPPFLPPPAKIVVDGSGCSTDTVKKTLDPKKNQHYDLYVK 64  
 Db 12 VKRLTYCACNLVKKPPFLPPPAKIVVDGSGCSTDTVKKTLDPKKNQHYDLYVK 71  
 QY 65 TDSITISVYNNHKKIHKQAGFLGCVLLSNASRLKDTGYRLDCLKLNPSFTDAVRGQ 124  
 Db 72 SDSVTISVYNNHKKIHKQAGFLGCVLLSNASRLKDTGYRLDCLKLNPSFTDAVRGQ 131  
 QY 125 IVVSLQTRDRIGTGGSVVDCKGLLEN-----GTVY----- 155  
 Db 132 IVVSLQSRDRIGTGGSVVDCKGLLEN-----GTVY----- 191

QY 156 --EDSGPGRPLSCFMEBPAPYTDSTGAAAGGNCRFVESPSODORLOARLNPPVRGSL 213  
 Db 192 ASEYSSGPRPLSCFVDENTPISGTNGATCG-----QSGDPRLAERARVYSGRHVM 242  
 QY 214 QTPQNPAGHQBELPEGEQRTTVQGOVYFLHTQGVSTWMDPRIPRLDINSVNDDELCP 273  
 Db 243 ---SRTHLHPDPLPEGEQRTTVQGOVYFLHTQGVSTWMDPRIPRLDINSVNDDELCP 298  
 QY 274 LPPGWEVRSVSGRIYFVDNNRTTQFTDPR---LHHIMNHCQLKEPSCPPLPSEGS 329  
 Db 299 LPPGWEIRNATGRVYFVDNNRTTQFTDPRISANLHLVANNQNLKQOQGOVY----S 354  
 QY 330 L---EDSELPAGRYERDLYOKLVTREHLSLQOPQAGCRLEVSREEIFEESYROIOMR 386  
 Db 355 LCPDDTECLTVPRYKRDLYOKLVTREHLSLQOPQAGCRLEVSREEIFEESYROIOMR 414  
 QY 387 PKDLKRLWVFRGEGGLDYGVAREMWYLICHEMLNPPYGLFOYSTNINWLOQNPSS 446  
 Db 415 PKDLKRLWVFRGEGGLDYGVAREMWYLICHEMLNPPYGLFOYSTNINWLOQNPSS 474  
 QY 447 INPDHLSYHFGRIYGLAVFHHYINGGFTVPPYKQLGKPIQLSDLESVDPELHSLV 506  
 Db 475 VNPEHLSYHFGRIYGLAVFHHYINGGFTVPPYKQLGKPIQLSDLESVDPELHSLV 534  
 QY 507 WLENDITPVLDHTCVENNAFGRILOHELKNGRNVPTENKKEYRLVYNNRPMRG 566  
 Db 535 WLENDITPVLDHTCVENNAFGRILOHELKNGRNVPTENKKEYRLVYNNRPMRG 594  
 QY 567 EAQFLAQKGFNELIPOHLKPPDQKELELIIGLDKIDLDWKSNTLKHCVADSNIVR 626  
 Db 595 EAQFLAQKGFNEVLIPOHLKPPDQKELELIIGLDKIDLDWKSNTLKHCVADSNIVR 654  
 QY 627 WFMQAVETDERRRRLQFVTSRVPLQGRKALQSGTGAAGPRLFTLHLIDANTDLP 686  
 Db 655 WFMQAVETDERRRRLQFVTSRVPLQGRKALQSGTGAAGPRLFTLHLIDANTDLP 711  
 QY 687 KATCFNRIDIPPESEYKYLTAVEETCGFAVE 723  
 Db 712 KATCFNRIDIPPESEYKYLTAVEETCGFAVE 748

RESULT 5  
 AAB31477  
 ID AAB31477 standard; protein; 748 AA.  
 AC AAB31477;  
 DT 20-APR-2001 (first entry)  
 DE Amino acid sequence of a human Smurf2 polypeptide.  
 DE Smurf1; Smurf2; Smad signal transduction; bone morphogenetic protein; BMP;  
 DE transforming growth factor-beta; human; TGF-beta; chondrogenesis;  
 DE osteogenesis; blood differentiation; cartilage formation; hair growth;  
 DE neural tube patterning; retinal development; heart induction;  
 DE morphogenesis; tooth formation; gamete formation.  
 OS Homo sapiens.  
 PN WO200071168-A2.  
 PD 21-DEC-2000.  
 PF 12-JUN-2000; 2000MO-US016250.  
 PR 11-JUN-1999; 99US-0138969P.  
 PA (UNY) UNIV NEW YORK STATE RES FOUND.  
 PA (HSCR-) HSC RES & DEV LP.  
 PI Thomsen GH, Wzana J;  
 DR WPI: 2001-071267/08.

DR N-PSDB; AAF24853.  
 XX  
 PT Novel isolated Smurf protein useful for inhibiting bone morphogenic  
 PT protein or tumor growth factor-beta activation pathway, for treating  
 PT cancer and to block osteogenesis, hair growth, tooth formation.  
 XX  
 PS Claim 10; Fig 12; 107pp; English.  
 XX  
 CC The present sequence represents a human Smurf2 polypeptide. The  
 CC specification also describes a Smurf1 polypeptide. Smurf polypeptides are  
 CC negative regulators of Smad signal transduction, and antagonists of bone  
 CC morphogenic protein (BMP) or transforming growth factor-beta (TGF-beta)  
 CC signaling pathway. Expression of Smurf1 in a cell is useful for  
 CC inhibiting a BMP or TGF-beta activation pathway in a cell. Smurf  
 CC polypeptides are useful for blocking chondrogenesis, osteogenesis, blood  
 CC differentiation, cartilage formation, neural tube patterning, retinal  
 CC development, heart induction and morphogenesis, hair growth, tooth  
 CC formation, gamete formation and a wide variety of tissue and organ  
 CC formation processes, and hinder the regeneration, growth, maintenance,  
 CC etc., of bone and other tissues that are dependent on the BMP pathway.  
 CC The polypeptide is useful for screening for various drugs and/or  
 CC antibodies that can either enhance the BMP pathway, or inhibit it  
 CC  
 XX  
 SQ Sequence 748 AA.  
 Query Match 77.3%; Score 3001; DB 4; Length 748;  
 Best Local Similarity 74.6%; Pred. No. 1,5e-285;  
 Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;  
 QY 5 IKTRLTLCAKMLAKKDFRLLPDPFAKIVYDGSQGSSTVTYKNTLDPKKNQHYDLYGK 64  
 DB 12 VKURLTLCAKMLAKKDFRLLPDPFAKIVYDGSQGSSTVTYKNTLDPKKNQHYDLYGK 71  
 QY 65 TDSITLTVNHNKHKIKKQAGAGLGCYVLLSNALSRLKDTGYOGLDLCINPSTDAVRQ 124  
 DB 72 SDSVTLTVNHNKHKIKKQAGAGLGCYVLLSNALSRLKDTGYOGLDLCINPSTDAVRQ 131  
 QY 125 IVVSLQTRDRIGTGGSVYDGRGLIENE-----GTIV----- 155  
 DB 132 IVVSLQTRDRIGTGGSVYDGRGLIENE-----GTIV----- 155  
 QY 156 --EDSGGRPLSCMEERPAFYDSTGAAGGACRFESPSODQRLQACRLNPDVAGSL 213  
 DB 192 ASBSSSSGRPLSCMEERPAFYDSTGAAGGACRFESPSODQRLQACRLNPDVAGSL 242  
 QY 214 QTPONRPHGHSPLPREGYBQRTTVQGVYFLHTQVGSVTHWDPRIKADLNSVNCDELGP 273  
 DB 243 ---SRTHLTPPDLIPREGYBQRTTVQGVYFLHTQVGSVTHWDPRIKADLNSVNCDELGP 298  
 QY 274 LPPGMEVRSVTSGRIFYVDHNHNRITTOFTDPR---LHINMHQCOLKEPQPLPLPSEGS 329  
 DB 299 LPPGMEVRSVTSGRIFYVDHNHNRITTOFTDPR---LHINMHQCOLKEPQPLPLPSEGS 354  
 QY 330 L---EDEELPQRYERDLVQKLKYLRLHLSLQOPQACHCRTEVSREBIFFBSTPQIMKMR 366  
 DB 355 LCPDTECLTVPRYKRDIVQKLKYLRLHLSLQOPQACHCRTEVSREBIFFBSTPQIMKMR 414  
 QY 387 PKDLKKRLMKVFRBEGLDYGVVAREMYLLCHEMLNPPYVLPQYSIDNTYMLQINPDS 446  
 DB 415 PKDLKKRLMKVFRBEGLDYGVVAREMYLLCHEMLNPPYVLPQYSIDNTYMLQINPDS 474  
 QY 447 INPDHLSYFHFVGRIMGLAVFHGYINGFTVPFYKQLGPKQLSDLESYDPELAKSLV 506  
 DB 475 VNPDLHSYFHFVGRIMGLAVFHGYINGFTVPFYKQLGPKQLSDLESYDPELAKSLV 534  
 QY 507 WLENDITPVLDHFTCVENHAFGRILCHELKNQGNVVEENKKEYRLVLVNRPMRGI 566  
 DB 535 WLENDITGVLDHFTCVENHAFGRILCHELKNQGNVVEENKKEYRLVLVNRPMRGI 594  
 QY 567 EAQFLALQKGFNELLPQHLKPFQDKLELLIIGGLDKIDLNDWKSNTLKHCVADSNIVR 626  
 DB 595 EAQFLALQKGFNELLPQHLKPFQDKLELLIIGGLDKIDLNDWKSNTLKHCVADSNIVR 654

QY 627 WFWQAVETPDERARRALLQVYTGSTRVPLQGFALQSGSTGAAGRPFTLHLDANTDNP 686  
 DB 655 WFWQAVETPDERARRALLQVYTGSTRVPLQGFALQSGSTGAAGRPFTLHLDANTDNP 711  
 QY 667 KAHTCFNRIDIPYSESYEKYEKLLTAVEETCGFAVE 723  
 DB 712 KAHTCFNRIDIPYSESYEKYEKLLTAVEETCGFAVE 748  
 RESULT 6  
 AAU19610  
 ID AAU19610 standard; protein; 804 AA.  
 XX  
 AC AAU19610;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Human diagnostic and therapeutic polypeptide (DITRP) #196.  
 XX  
 KW Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
 KW cell proliferative disorder; Crohn's disease; lymphoma; leukemia;  
 KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
 KW respiratory disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200162927-A2.  
 PD 30-AUG-2001.  
 XX  
 XX 21-FEB-2001; 2001WO-US006059.  
 XX  
 PR 24-FEB-2000; 2000US-0184693P.  
 PR 24-FEB-2000; 2000US-0184697P.  
 PR 24-FEB-2000; 2000US-0184698P.  
 PR 24-FEB-2000; 2000US-0184768P.  
 PR 24-FEB-2000; 2000US-0184769P.  
 PR 24-FEB-2000; 2000US-0184770P.  
 PR 24-FEB-2000; 2000US-0184771P.  
 PR 24-FEB-2000; 2000US-0184772P.  
 PR 24-FEB-2000; 2000US-0184773P.  
 PR 24-FEB-2000; 2000US-0184774P.  
 PR 24-FEB-2000; 2000US-0184776P.  
 PR 24-FEB-2000; 2000US-0184777P.  
 PR 24-FEB-2000; 2000US-0184797P.  
 PR 24-FEB-2000; 2000US-0184813P.  
 PR 24-FEB-2000; 2000US-0184817P.  
 PR 24-FEB-2000; 2000US-0184841P.  
 PR 24-FEB-2000; 2000US-0185213P.  
 PR 24-FEB-2000; 2000US-0185216P.  
 PR 12-MAY-2000; 2000US-0203785P.  
 PR 15-MAY-2000; 2000US-0204226P.  
 PR 15-MAY-2000; 2000US-0204525P.  
 PR 16-MAY-2000; 2000US-0204821P.  
 PR 16-MAY-2000; 2000US-0204908P.  
 PR 16-MAY-2000; 2000US-0205232P.  
 PR 17-MAY-2000; 2000US-0204815P.  
 PR 17-MAY-2000; 2000US-0204833P.  
 PR 17-MAY-2000; 2000US-0205221P.  
 PR 17-MAY-2000; 2000US-0205285P.  
 PR 17-MAY-2000; 2000US-0205286P.  
 PR 17-MAY-2000; 2000US-0205287P.  
 PR 17-MAY-2000; 2000US-0205323P.  
 PR 17-MAY-2000; 2000US-0205324P.  
 XX  
 PA (INCYTE) INCYTE GENOMICS INC.  
 XX  
 PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC,  
 PI Chen A, D'sa SA, Amey S, Dahl CR, Dam TC, Daniels SE, Dufour GE,  
 PI Flores V, Fong WT, Greenwalt LB, Hillman JL, Jones AJ, Liu TF,  
 PI Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Datto A,  
 PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W,  
 PI Cohen HT, Hodgson DM, Lincoln SE, Jackson S;

XX WPI; 2001-502667/55.  
 DR N-PSDB; AAS31181.  
 XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
 PT enzymes, hormones and receptors, useful in diagnostics and therapeutics.  
 PS Claim 27; Page 512-514; 522pp; English.  
 CC The invention relates to polynucleotides (I) encoding diagnostic and  
 CC therapeutic (DTHP) polypeptides (II), which include e.g. enzymes, and  
 CC proteins involved in growth and development and receptors. (I) and (II)  
 CC may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate DTHP expression. For example, (I) and (II)  
 CC may be used to treat disorders associated with decreased polypeptide  
 CC expression by rectifying mutations or deletions in a patient's genome,  
 CC that affect the activity of the DTHPs, by expressing inactive proteins  
 CC or supplementing the patient's own production of them. (I) and (II) may  
 CC be used to treat diseases, for example, cell proliferative disorder,  
 CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
 CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
 CC (I) may be used to produce the DTHPs, by inserting the nucleic acids  
 CC into a host cell and culturing the cell to express the protein. (I) and  
 CC its complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantify the presence of similar nucleic acids in  
 CC samples, and therefore which patients may be in need of restorative  
 CC therapy. (II) may also be used as antigens in the production of  
 CC antibodies against DTHPs and in assays to identify modulators of DTHP  
 CC expression and activity. The anti-DTHP antibodies and antagonists may  
 CC also be used to down regulate expression and activity. The anti-DTHP  
 CC antibodies may also be used as diagnostic agents for detecting the  
 CC presence of DTHPs in samples (e.g. by enzyme linked immunosorbent assay  
 CC (ELISA)). AAU9415-AAU9625 represent human diagnostic and therapeutic  
 CC (DTHP) polypeptides of the invention  
 XX  
 XX Sequence 804 AA;  
 SQ  
 Query Match 77.1%; Score 2996; DB 4; Length 804;  
 Best Local Similarity 74.5%; Pred. No. 5.4e-25;  
 Matches 564; Conservative 63; Mismatches 72; Indels 58; Gaps 8;  
 QY 5 IKIRITVLCANLAKKDFRLDPFAKIVVDGSGQCHSTDTVKTLPKXNQHLYVVK 64  
 DB 68 VMLHLTVLCANLTVKDFRLLDPFAKIVVDGSGQCHSTDTVKTLPKXNQHLYVIGK 127  
 QY 65 TDSITISVWNHKKTHKXGAGFLGCVRLSNALSLMDTGYQRDLCKLPSDIDANRQ 124  
 DB 128 SDSVITISVWNHKKTHKXGAGFLGCVRLSNALSLMDTGYQRDLCKLPSDIDANRQ 167  
 QY 125 IVSLIQTRDRIGTGGSVVDGCRGLDNE-----GTVY----- 155  
 DB 188 IVSLIQSRDRIGTGGSVVDGCRGLDNE-----GTVY----- 155  
 QY 156 --EDSGRRLSCMEPEAPYTDSTGAAGGNCRFVESPQDORLQAQRLNPDVAGSL 213  
 DB 248 ASEYSSPRLSCVDEVTPISTGNAGACG-----OSSDRLEARRVRSQRHNAYM 298  
 QY 214 QTPNRRPHGQSPPELPEGEQRTVQGVYFLHTQTGVSTWHDRIIPRDLNSVNCDELGP 273  
 DB 299 ----SRTLHTRPPLPESYEQRITQOGGVYFLHTQTGVSTWHDRIIPRDLNSVNCDELGP 354  
 QY 274 LPPGMEVASTVSGRIYFDHNNRTTQTDPR---LHIMNHQCLKEPSQPLPLPESGS 329  
 DB 355 LPPGMEVASTVSGRIYFDHNNRTTQTDPR---LHIMNHQCLKEPSQPLPLPESGS 329  
 QY 330 L--EDELPAQRYERDLYVOKLKYLRHSLQQAQCHRIEVSREIPEESROIMKMR 366  
 DB 411 LCPDTECLIVPRKRDLYVOKLKYLRHSLQQAQCHRIEVSREIPEESROIMKMR 470  
 QY 387 FKDIKKRLMVKREGEGLDYGVARLWYLLCHEMLNPPYGLFQYSTDNITMQLINPDS 446  
 DB 471 FKDIKKRLMVKREGEGLDYGVARLWYLLCHEMLNPPYGLFQYSTDNITMQLINPDS 530

QY 447 INPDHLSYFHFVGRINGLAVFHGTYNGGFVTPVPEYKQLLGKPIQLSDLESVDPELHKSIV 506  
 DB 531 VNPETHLSYFHFVGRINGLAVFHGTYNGGFVTPVPEYKQLLGKPIQLSDLESVDPELHKSIV 590  
 QY 507 WLENDITPVLDHTPCVCENNAFGRILOHELKPNGRNVPYTEENKKEYVLYVNWRFMRGI 566  
 DB 591 WLENDITGVLDHTPCVCENNAFGRILOHELKPNGRNVPYTEENKKEYVLYVNWRFMRGI 650  
 QY 567 EAQFLAQGFENLIDQHLKPKDPCKELIIIGSLDKIDINMKSNTRIKHVAASNTVR 626  
 DB 651 EAQFLAQGFENLIDQHLKPKDPCKELIIIGSLDKIDINMKSNTRIKHVAASNTVR 710  
 QY 627 WFWQAVETDEERRARLLQFTVGTSTVPLQGFALQSGTGAAGPRLFTIHLIDANTDNL 686  
 DB 711 WFWQAVETDEERRARLLQFTVGTSTVPLQGFALQSGTGAAGPRLFTIHLIDANTDNL 767  
 QY 687 KAHTCFNRIDIPYSEYKELKLTAVETGFAVE 723  
 DB 768 KAHTCFNRIDIPYSEYKELKLTAVETGFAVE 804  
 RESULT 7  
 ID AAU87301 standard; protein; 514 AA.  
 AC AAU87301;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Novel central nervous system protein #211.  
 KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
 KW hyperproliferative disorder; neoplasia; cardiovascular disorder;  
 KW cardiac arrest; cerebrovascular disorder; ischemia; angioneurosis;  
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
 KW endocrine disorder; diabetes; cancer; leukemia; neovascularization;  
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
 KW myocardial infarction; wound healing; cell proliferation; skin aging;  
 KW food additive; food preservative; gene therapy.  
 KW Homo sapiens.  
 XX  
 XX MO200155318-A2.  
 PD 02-AUG-2001.  
 XX  
 XX 17-JAN-2001; 2001WO-US001332.  
 PR 31-JAN-2000; 2000US-0179065P.  
 PR 04-FEB-2000; 2000US-0180628P.  
 PR 24-FEB-2000; 2000US-0184664P.  
 PR 02-MAR-2000; 2000US-0186350P.  
 PR 16-MAR-2000; 2000US-0189874P.  
 PR 17-MAR-2000; 2000US-0190076P.  
 PR 18-APR-2000; 2000US-0198123P.  
 PR 19-MAY-2000; 2000US-0205515P.  
 PR 07-JUN-2000; 2000US-0209467P.  
 PR 28-JUN-2000; 2000US-0218866P.  
 PR 30-JUN-2000; 2000US-0215135P.  
 PR 07-JUL-2000; 2000US-0216647P.  
 PR 11-JUL-2000; 2000US-0216880P.  
 PR 11-JUL-2000; 2000US-0217487P.  
 PR 11-JUL-2000; 2000US-0217496P.  
 PR 14-JUL-2000; 2000US-0218295P.  
 PR 26-JUL-2000; 2000US-0220963P.  
 PR 26-JUL-2000; 2000US-0220964P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0224519P.  
 PR 14-AUG-2000; 2000US-0225213P.  
 PR 14-AUG-2000; 2000US-0225214P.  
 PR 14-AUG-2000; 2000US-0225266P.



CC maintain organs before transplantation, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 57.6%; Score 2236; DB 4; Length 514;  
Best Local Similarity 80.7%; Pred. No. 1,8e-210;  
Matches 414; Conservative 43; Mismatches 42; Indels 14; Gaps 4;

QY 218 NRPQHOSPELPEGEORTTVOGVYFLHTGTSTWMDPRI PRDLNSVNCDELGPPLPG 277  
DB 9 SETHHTPPDPEGEORRTTQGGYFLHTGTSTWMDPRVRLNSINCEELGPPLPG 68  
QY 278 WVRSTVSGRIYFVDHNNRTTQFTDPR---LHHMNNCCOLKPSQPLPSPSGSL--- 330  
DB 69 WEIRNTATGRVYFVDHNNRTTQFTDPRISANLHVLNQNOLKQOQOQV---SLCPD 124  
QY 331 EDEELPAQYERDVLQKLVRLHSLQOPAGHCRLEVSREELFEESYRQIMKRPDL 390  
DB 125 DTECLTVPYKRDVLQKLVRLHSLQOPAGHCRLEVSREELFEESYRQIMKRPDL 184  
QY 331 KRLLWVFRGREGLDYGVAREMLYLCHENLNPYGLFOYSTNTIYMLQINPSSINPD 450  
DB 135 WKRLMTKRGEEGLDYGVAREMLYLSHEMLNPYGLFOYSRDIYTLQINPSSAVNPE 244  
QY 451 HLSTYFHVGRINGLAVFHGHYINGFTVPFYKOLLGKEIQLSDESYDPELHKSILWILE 510  
DB 245 HLSTYFHVGRINGLAVFHGHYIDGGFTLPFYKOLLGKSTILDVHVLDPDLHSLWILE 304  
QY 511 NDITFVLDHTECVENNAFGRLLQHLHLEKNGRNVPTENKKEYRLVLYNMFNGRTIAOF 570  
DB 305 NDITFVLDHTECVENNAFGRLLQHLHLEKNGKSIPEVNEKKEYRLVLYNMFNGRTIAOF 364  
QY 571 LALQGFNELIPHLLKFPDQKELELIGLKDIDLNDWKSNTLKHCVASDNIYVRFWQ 630  
DB 365 LALQGFNEVLPQHLKTFDEKELELIGLKDIDLNDWKSNTLKHCVASDNIYVRFWQ 424  
QY 631 AVETDERRRRLLOFTGSTRVPLQGFKALQSTGAAPRLFTIHLIDANTDNLPRAT 690  
DB 425 AVETDERRRRLLOFTGSTRVPLQGFKALQSTGAAPRLFTIHLIDANTDNLPRAT 481  
QY 691 CFNRIDIPYSEYKLEKLTAVEETGGFAVE 723  
DB 482 CFNRIDIPYSEYKLEKLTAVEETGGFAVE 514

RESULT 8  
ABB61120  
ID ABB61120 standard; protein, 1035 AA.

XX AC ABB61120;  
XX DT 26-MAR-2002 (first entry)  
XX DE Drosophila melanogaster polypeptide SEQ ID NO 10152.  
XX KW Drosophila; developmental biology; cell signalling; insecticide;  
XX OS pharmaceutical.  
XX OS Drosophila melanogaster.  
XX PN WO200171042-A2.  
XX PD 27-SEP-2001.  
XX PF 23-MAR-2001; 2001WO-US009231.  
XX PR 23-MAR-2000; 2000US-0191637P.  
XX PR 11-JUL-2000; 2000US-00614150.  
XX PA (PEKE) PE CORP NY.  
XX Venter JC, Adams M, Li FWD, Myers EM;

XX WPI, 2001-656860/75.  
XX DR N-PSDB; ABL052233.  
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
PS Disclosure; SEQ ID NO 10152; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC ABBS78072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WPIO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 1035 AA;

Query Match 53.3%; Score 2069; DB 4; Length 1035;  
Best Local Similarity 41.2%; Pred. No. 1.7e-193;  
Matches 435; Conservative 104; Mismatches 145; Indels 372; Gaps 15;

QY 6 KIRLTYLCAKNTAKDPRLLPPEFAKIYDGGQCHSDTYKNTIDPRKMGHYLYVQKT 65  
DB 14 KVRITTLCAKNTAKDPRLLPPEFAKIYDGGQCHSDTYKNTIDPRKMGHYLYVQKT 73  
QY 66 DSITTSVNMHKKIKKQAGFLGCVRLSNASISRLKDGORYOLDCKINPBDTAVRQOI 125  
DB 74 DAITTVNMQRKIRH--GSGFLGCVRIAPFNIGSLKAGFGRLDGKSLSPDDDELVRQOI 131  
QY 126 VVSLQTRDRIGTG-----GSVYDCRGLENE-----GVYV----- 155  
DB 132 IISLSKQFSSGNPLAIVGSGDVGPSDDSDSSDSLPGEBSERRTDNGRYYVNHATK 191  
QY 156 ----- 155  
DB 192 STQWRPAPQGVGSSHATSPQQRNTNNGSGDRQAPAGTRSTCTNLMNNGHRSRL 251  
QY 156 -----EDSGPRPLSCFMEEPAYDTSTGAAGGNCREVE----- 191  
DB 252 SVTASDERRHSTELISVCKENTSPPTVSA-TTPGKKTSSSSSSAGG--RTLEGQPT 308  
QY 192 -----SPSQDQRLQARLNPDPVRSLOTPQNRPHGQSP----- 227  
DB 309 NEPATPTSTTSASVRLHSND--NHVKTPKQITNGHAPBSPTSPGQOQYVNGNAGNG 366  
QY 228 ----- 227  
DB 367 STSGNGSGQAQPOGSANGWTQEDAAATTSPSTTSPPRHSQSPPTNISPPASVTSAN 426  
QY 228 ----- 227  
DB 427 GNVHSPNANSTPAGSGGSRSYTAATPGQSRSSRQOGESSSTRRRSGRTNGGTSG 486  
QY 228 -----LPEGEORTTVOGVYFLHTGTSTWMDPRI 259  
DB 487 GGGGGSGQRYASAAIAAANGARFLLDPRGYEMRTTQGGYVFIHPIGVSTWMDPRI 546  
QY 260 PRDLNS--VNCDELGPPLPGWEVRSYSGRIYFVDHNNRTTQFTDPRLHHIM----- 309  
DB 547 PRDPTQHLTLDAIGPLPSGWEORKTASGRVYFVDHNNRTTQFTDPRLSGILQMRGT 606  
QY 310 ---NHQOQLKPSQPLPSPS-----EG 328  
DB 607 VPPTSANAGTPAPPSATPAPPSAAAAPPQATPASNATPTTLTTTNPPHRIYVDDLPQG 666  
QY 329 SLEDEELPAQYERDVLQKLVRLHSLQOPAGHCRLEVSREELFEESYRQIMKRPK 388

```

Db      667 LLEGADL-LPKYARDLVGKRLALRTELQTMQPSGCHCRLEVSNNRFEESYRLIMYRAK 725
Qy      389 DLKKRLMVKRGGEGLDYGVAREMYLTLCHENLNTYYGLFQYSTNTIYMLQINPSSIN 448
Db      726 DMKRLMVKRGGEGLDYGVAREMYLTLCHENLNTYYGLFQYSTNTIYMLQINPSSIN 785
Qy      449 PDHLSYFHFVGRIMGLAVFHGYINGFVTFYFKQILKPIQLSDLESYDEPLHKSILWMI 508
Db      786 PDHLSYFHF-----LLNKRTITGDLBGVDPDLHSLTMM 819
Qy      509 LENDIPVLDHTFCVENHAFGLQLHLEKNGNNVPTENKKEYRLIYNNRMFCIGIA 568
Db      820 LENSISGIISTSVENNSFGALVHEHLKPGASIPTEENKREYKLYLVNRYFRMGIRQ 879
Qy      569 QFIALQKGFNELLPOHLKRPDOKELIIGLGDKIDLNDMKSNTRLKQVADSNIVRMF 628
Db      880 QFIALQKGFCELLPSHLRPFDERELVIGISIDVNDNRNTLTKCTNETTQVLMF 939
Qy      629 WQAVEFDEERRARLLQFTVGTSTVPLOQFKALQSGTGAAPRLFTIHL-IDANTDNLK 687
Db      940 WQVYESSEMRRARLLQFTVGTSSRVPLOQFKALQSGTGAAPRLFTIHLTADVPQNLK 999
Qy      688 AHTCFNRIIDIPYSEYKLYEKLITAVEETCGFAVE 723
Db      1000 AHTCFNRIIDIPYETVQLDCLTLQAVEETCGFAVE 1035

RESULT 9
AAM78877
ID      AAM78877 standard; protein; 335 AA.
XX
AC      AAM78877;
XX
XX      06-NOV-2001 (first entry)
XX
DE      Human protein SEQ ID NO 1539.
XX
XX      Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX      vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX      tissue growth factor; immunomodulatory; cancer; leukaemia;
XX      nervous system disorder; arthritis; inflammation.
XX
OS      Homo sapiens.
XX
PN      W02001:57190-A2.
XX
PD      09-AUG-2001.
XX
PF      05-FEB-2001; 2001WO-US004098.
XX
XX      03-FEB-2000; 2000US-00496814.
XX      27-APR-2000; 2000US-0050875.
XX      20-JUN-2000; 2000US-00598075.
XX      19-JUL-2000; 2000US-00620325.
XX      01-SEP-2000; 2000US-00654936.
XX      15-SEP-2000; 2000US-00663561.
XX      20-OCT-2000; 2000US-00693325.
XX      30-NOV-2000; 2000US-00728422.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
XX      Ma Y, Zhao QA, Wang J, Zhang J, Ren F, Chen R, Wang ZW,
XX      Xue AJ, Yang Y, Wejthman T, Goodrich R;
XX      WPI; 2001-476283/51.
XX      DR      N-PSDB; AAK52010.
XX      Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX      PT      in diagnosis and gene therapy.
XX      PS      Claim 20; Page 3847; 6221pp; English.
XX

```

```

CC      The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC      encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC      cytokine, cell proliferation or cell differentiation or which may induce
CC      production of other cytokines in other cell populations. The
CC      polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC      peptide therapy. The polypeptides have various cytokine-like activities,
CC      e.g. stem cell growth factor activity, haematopoiesis regulating
CC      activity, tissue growth factor activity, immunomodulatory activity and
CC      activin/inhibin activity and may be useful in the diagnosis and/or
CC      treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC      inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC      (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC      sequence listing were missing at the time of publication
XX
SQ      Sequence 335 AA;
XX
Query Match      41.6%; Score 1614.5; DB 4; Length 335;
Best Local Similarity 86.2%; Pred. No. 1.5e-149;
Matches 306; Conservative 4; Mismatches 20; Indels 25; Gaps 2;

Qy      168 MEPPAPYTDSTGAAGAGNCRFVESPSQDQRLQARLNPVRSGLQTPQNRPHGHSPE 227
Db      1 MEPPAPYTDSTGAAGAGNCRFVESPSQDQRLQARLNPVRSGLQTPQNRPHGHSPE 60
Qy      228 LPEGYEORTVVOGVYFHTQGVSTWHPRIIPRLNSVNDDELGLPLPGWEVSTVSGR 287
Db      61 LPEGYEORTVVOGVYFHTQGVSTWHPRIIPRLNSVNDDELGLPLPGWEVSTVSGR 120
Qy      288 IYFVDHNRRTQFTDPRLHHIMHQCQKPEPQPLPSEGSLEDEHLPQRYERDLYQK 347
Db      121 IYFVDHNRRTQFTDPRLHHIMHQCQKPEPQPLPSEGSLEDEHLPQRYERDLYQK 180
Qy      348 LKVLRLHLSLQOPQAGHCRIEVSREIPEESYRQIMKRPKDLKRLMVKRGGEGLDYG 407
Db      181 LKVLRLHLSLQOPQAGHCRIEVSREIPEESYRQIMKRPKDLKRLMVKRGGEGLDYG 240
Qy      408 GVAEEMLYLCHENLNTYYGLFQYSTNTIYMLQINPSSINPDHLSYFHFVGRIMGLAVF 467
Db      241 GVAEEMLYLCHENLNTYYGLFQYSTNTIYMLQINPSSINPDHLSYFHFVGRIMGLAVF 298
Qy      468 HGHYINGFTVPFYKQLGKPIQLSDLESYDEPLHKSILWILENDITVLDHTFC 522
Db      299 -----CVPWTLHQRG-----LHSAALLQAAAGRAHPLRSGIC 330

RESULT 10
AAM13385
ID      AAM13385 standard; protein; 766 AA.
XX
AC      AAM13385;
XX
XX      10-UTL-1997 (first entry)
XX
DE      Human protein ubiquitin ligase pub2.
XX
XX      Protein ubiquitin ligase; pub2; cdc25 phosphatase; CDK kinase; p53;
XX      cell cycle; transgenic animal.
XX
OS      Homo sapiens.
XX
PN      W09712962-A1.
XX
PD      10-APR-1997.
XX
PF      04-OCT-1996; 96WO-US015930.
XX      04-OCT-1995; 95US-00539205.
XX
PA      (COLD-) COLD SPRING HARBOR LAB.
XX      Beach D, Caligiuri M, Nefeky B;
XX      WPI; 1997-226206/20.
XX

```

DR N-PSDB; AAT47041.  
 XX Human and Saccharomyces pombe protein ubiquitin ligase (s) - involved in  
 PT cdc35 phosphatase and p53 ubiquitination, and regulate cell growth and  
 PT proliferation.  
 XX  
 XX Claim 1; Page 80-84; 108pp; English.  
 XX  
 CC Human protein ubiquitin ligases pub1 (AAW13384), pub2 (AAW13385) and pub3  
 CC (AAW13386) are homologues of fission yeast pub1 (AAW13387) and were  
 CC identified from cDNA clones (AAT47040-42) obt. e.g. from a keratinocyte  
 CC cDNA library. Pub polypeptides can be produced in transfected host cells.  
 CC They can control the steady state level of cdc35 phosphatase, the degree  
 CC of CDK kinase (e.g. cdc2) dephosphorylation and the steady state level of  
 CC p53 (controlling the degree of cell cycle regulation of p53). They can  
 CC regulate cell or tissue differentiation, or cell growth or proliferation  
 CC by affecting other proteins, can be a specific (ant)agonist of wild-type  
 CC protein function and may be used as immunogens to elicit a specific  
 CC immune residue  
 CC  
 SQ Sequence 766 AA;  
 Query Match 37.6%; Score 1461; DB 2; Length 766;  
 Best Local Similarity 40.8%; Pred. No. 8.2e-134;  
 Matches 326; Conservative 118; Mismatches 236; Indels 120; Gaps 20;  
 QY 4 SIKIKLITLCAKNAKPPFRLPPFAKIVDGGSCCHSTPTVKTLDPKNQHLYVG 63  
 DB 7 SRRIVIIIVAADGLYKRVFRFPDPFAVLTVDGE-QTHTTAKKTLNPNYNETFEVAVT 65  
 QY 64 KTDSTIISVMHKKIKKQAGFLGCVALLSNISRLKDTGYQRL-DLCLNTNSPDYAV 121  
 DB 66 DNSTAIQVFDQKF-KKKGGGFLGVINLRGVDVLDLAGDEMLICDLKLN-EMTVV 122  
 QY 122 RQIVVSLQ-----TRDRIGTGGSVYDCKGLLENQTYED 157  
 DB 123 HKIITINISTAQTLQVPSAASGARTRQTSITNDPSSSSSVSRNPASRAGSPTRD 182  
 QY 158 SGPQ-----RPLSCFMEB---PAPY---TDSYG-----AAA 182  
 DB 183 NAPAASPASSERTSSFEQYGRLPFGMEKTDNLGTYVDNHTSTWRNLSVA 242  
 QY 183 GGCNCRFVESPS-----ODRLQAGRLNPDVRSGLQTPQNRPHGOSPELPEGEYQ 234  
 DB 243 GAAAEHLSSASNAVTEGVQSSSNARTEASVLTSMNTAG-----SGELLPMGEQ 296  
 QY 225 RTTVGGQVYELHTQGVSTWHDPRIRPDLNSV-----CDELGPFGMEVYST 283  
 DB 297 RTTPGRFVVDANTRTTTWDPDRQYTRSYGPNNAATIQQPVSQLGFLPSGEMERLT 356  
 QY 284 VSGRIYVDHNNRTQFTDPRLHIMNHQCLKEPSQPLPSPSGSEDEELPAQRYERD 343  
 DB 357 NTARVYFVDHNTKTTMDPR-----LPS--SL-DQVVP--QYKRD 392  
 QY 344 LVQKIKVLRHELSLQOPQAGHCRLEVSREIPEBSYROIIMCRKDLKRLMYFRBEG 403  
 DB 393 FRRKLIYLSQPAL-HPLPGQCHKVRNHLFEOSVAIMQSAITDLKRLIMIFDEBDG 451  
 QY 404 LDYGVAREWLYLCHENLNPPYGLFOYSTDNIVYLQINPSSINPDLSTFHFVGRING 463  
 DB 452 LDYGLSEVFFLHSHENPNPYCLFEYSVDNTLQINPNSGINPELNAFKRIGVIG 511  
 QY 464 IAVFHGHTINGFTVFFYKOLKPIQLSDLESVDPELHKLVLWLENDITPVIDHTPCV 523  
 DB 512 LAIFRRFVDVAFVVSFYKMTLOKKVTLQDMESMDAEYRSVLWINDITGVLDLTFVS 571  
 QY 524 EHNAAGRILOHLEKNGSNVPTTEENKKEYRVLVYNNRPFMGIAOPLAOKGNEILPQ 593  
 DB 572 EDNCGEYVITDKNGNRIEYTEENKREYVDLVTV--IQRIEQRPAHFEHSGSLIPQ 630  
 QY 584 HLKFPDQKLELLITIGLDKIDLNDWKSNTLKEICVADSNIVRMFWOAVETFEDEEARL 643  
 DB 631 ELINVFDERRELLIGISIEIDMEDWKHGDYRSYSENDQIITKPFWEIMDEWMSKKSRL 690

QY 644 LQFVGTSTVPLQGFKALQSGTAGAPRLFTLIDANTDLPKATGCRNIDIPPEBSY 703  
 DB 691 LQFTTSTSLPVGFFDLQSGD---GPRKFTLEKA-GERNKLPKATGCRNLDLPYNSX 746  
 QY 704 EKLYEKULIATVEETGFAVE 723  
 DB 747 KDLDRKLSIAVEETIGFQGE 766  
 RESULT 11  
 ID ABP73459  
 XX ABP73459 standard; protein; 832 AA.  
 ID ABP73459  
 AC ABP73459;  
 DT 30-JAN-2003 (first entry)  
 DE Candida albicans essential protein SEQ ID NO 7296.  
 XX  
 XX Fungus; yeast; tetracyclin; promoter; GRACE strain; biosynthesis;  
 KM signal transduction; DNA replication; cell division; growth;  
 KM proliferation; Candida albicans; fungicide; antifungal.  
 XX  
 OS Candida albicans.  
 XX  
 PN W0200253728-A2.  
 XX  
 PD 11-JUL-2002.  
 XX  
 PF 26-DEC-2001; 2001WO-US049486.  
 XX  
 PR 29-DEC-2000; 2000US-0259128P.  
 PR 20-FEB-2001; 2001US-0079202A.  
 PR 22-AUG-2001; 2001US-0314050P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Roemer T, Jiang B, Boone C, Bussey H, Ohlsen KL;  
 XX  
 DR WPI, 2002-566694/60.  
 DR N-PSDB; AB232009.  
 XX  
 PT Constructing strains for identifying gene products as effective targets  
 PT for therapeutic intervention, by inactivating in the strain one allele of  
 PT a gene and placing other allele of the gene under conditional expression.  
 XX  
 PS Claim 44; SEQ ID NO 7296; 167bp + Sequence Listing; English.  
 XX  
 CC The invention relates to constructing (M1) a strain of diploid fungal  
 CC cells in which both alleles of a gene are modified, comprising modifying  
 CC one allele by insertion or replacement by a cassette having an  
 CC expressible selectable marker and modifying other allele by  
 CC recombination of a promoter replacement fragment with a heterologous  
 CC promoter, so that expression of the second allele is regulated by the  
 CC promoter. (M1) is useful for constructing a strain of diploid fungal  
 CC cells having both alleles modified of a gene are modified. The diploid fungal  
 CC cells in which both alleles modified are useful for identifying a gene that  
 CC is essential to the survival or growth of a fungus, a gene that  
 CC contributes to the virulence and/or pathogenicity of a fungus, a gene  
 CC that contributes to the resistance of a diploid fungus to an antifungal  
 CC agent, an antifungal agent that inhibits the growth of a diploid fungus  
 CC and for identifying a therapeutic agent for treatment of a mammalian  
 CC disease. (M1) is useful for identifying a compound which modulates the  
 CC activity of a gene product, preferably enzymatic activity, carbon  
 CC compound catabolism, biosynthetic, transporter, transcriptional,  
 CC translational, signal transduction, DNA replication and cell division  
 CC activity. The method is useful for identifying a compound having the  
 CC ability to inhibit growth or proliferation of C. albicans cells and for  
 CC treating infection by C. albicans. The present sequence is that of an  
 CC essential Candida albicans protein used in the method of the invention.  
 CC Note: The sequence data for this patent is not represented in the printed  
 CC specification but is based on sequence information supplied to Derwent by

```
CC the European Patent Office
XX
SQ Sequence 832 AA;

Query Match
  37.0%; Score 1436.5; DB 5; Length 832;
Best Local Similarity 38.4%; Pred. No. 2.5e-131;
Matches 326; Conservative 124; Mismatches 233; Indels 167; Gaps 21;

QY 7 IRLTVLCAKNLAKOFFRLDPFAKIVVDGSGQCHSTDTVKTLDPKWNQHYDLVVGKTD 66
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 17 INKVAAESLYKRDVFRQDPDFAVLTVDGS-QTKTITAKTLNPNYNETFN-FOAKED 74
QY 67 ST-TTSVWNHKKIHKQAGFLGCVRL--LSNAISRLKDTGVQRL--DLCKLNPSDIDAVR 122
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 75 SILVQVFOQKF-KKQDQFLGVINVLGVIDLSLNSSEETITRDLLKKN--ENLAVS 131
QY 123 GOIVVSLQ-----TRD-----RIGTGGSVVDCRGLLE 149
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 132 GXIIVISHNRNSGGVTTATTGTGASSNNIATITSGVNNLRIGSATTTANSTAQAS 191
QY 150 NEGTV-YEDSGPGR---PLSCFWEPPAPYDSTGAAGGNCRFVESQDORLOAQLRL 205
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 192 SDATVANGSGPTSLPIGQGHPEATATPGGAAGAASQYSSFEDQYGRLPFGWER 251
QY 206 NPDVRG-----SLQTPQNRPHGHQSP----- 226
DQ 252 RTDNFGRTYVDHNSRTTWTQREPALHQSETERGQQRQSETEAERQHRGRTLPGSGVSP 311
QY 227 -----ELPEGYEORTVQOV 242
DQ 312 LPTGSGNITSNGVTNAGANTPNPAAAVSMAASGATTSLGLSLPGSGWQRFTTEGRP 371
QY 243 YELHTQTGVSTWHDPRIPDLNLSVN-----CDELGPLPPGWEVSTVSGRIYFVDH 293
DQ 372 YFVDENRTTWTVDPRQYIITFGNTTIOQPVSQLGPLSGWEMRLTNTARYFVDH 431
QY 294 NNRTOFTDPRLHIMNHOCQLKEPSQPLPSEGSLEDELPQRYERDLVQKLKVRH 353
DQ 432 NTKTTTWDPR-----LPS--SL-DQNVF--QYRDFRFRKVIYFRS 467
QY 354 ELSLOQPOAGHCRIEVSREEIFEESVQIMKMRPKDLKRLMKVFRGEGLDYGVAREW 413
DQ 468 QVAL-RILPQCHIKVRRDHIFEDSYQELMRQTPEDLKKRLMKPDGEGLDYGVAREW 526
QY 414 LYLICHEMLNPPYGLFYQSTDNIMQLNPDSINPDHLSYPHFVGRINGLAVPHGHYIN 473
DQ 527 FELLSDHMFNPFYCLFEYSSHDNYTLQINPNSGINPEHLNFKFGRVVLGVFHRFLD 586
QY 474 GGFVPPYKQLLGKPTQLSDLESVDPELHKSILVWILENDITVLDHTFCVHNATGRILO 533
DQ 587 AFFVGALYKQMLHKVYLQDMGVDABFYRSUKWILDNDITGILDITSAEESFGEIVE 646
QY 534 HELKPNRNVFVTEENKKEYVLYNWRPMRGIEAQFLALQKGFNELIPQHLLKPFDOKE 593
DQ 647 VDLKFGRIEIVTEENKHEYVELITWRLSKRVEQKAFIDGFNELIPQELVNVFDERE 706
QY 594 LELITGLDKIDLNDWKSNTLKHCVADSNIVRWQAVETDEBRRLALLOFTVGTSTV 653
DQ 707 LELLIGLAEIDCEDWKKHTYRGQENDQVQWFWKNCINWDSDQKARLLOFTTGTGRI 766
QY 654 PLQGFKALOGSTGAAGPRFTTHLIDANTDNLPKAHTCFNRIDIPPYSEYKLYSKLLTA 713
DQ 767 PVNGFKDLQSD---GPRFTTEKA-GEANQPKSHTCFNRVDLPPYTDYSLKQKLLTA 822
QY 714 VEETCGFAVE 723
DQ 823 VEETVGFQGE 832

RESULT 12
ID ABJ26104
XX standard; protein; 869 AA.

AC ABJ26104;
XX
DT 16-APR-2003 (first entry)
XX
DE Aspergillus fumigatus essential gene protein #762.
XX
KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX cancer; contamination; biofilm; antibody; immune response.
XX
OS Aspergillus fumigatus.
XX
PN WO200286090-A2.
XX
PD 31-OCT-2002.
XX
EF 23-APR-2002; 2002WO-US013142.
XX
PR 23-APR-2001; 2001US-0285697P.
XX 27-APR-2001; 2001US-0287066P.
PR 05-JUN-2001; 2001US-0295890P.
PR 09-JUL-2001; 2001US-0303899P.
PR 31-AUG-2001; 2001US-0316362P.
XX
FA (ELIT-) ELITRA PHARM INC.
XX
PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.
XX
DR
XX
PT New purified or isolated nucleic acids of essential genes of Aspergillus
PT fumigatus, useful for treating or preventing infections by A. fumigatus,
PT for treating a non-infectious disease in a subject e.g. cancer.
XX
PS Disclosure; Page; 175pp; English.
XX
CC The invention relates to novel purified or isolated nucleic acids of
CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
CC the invention are used to treat or prevent infections by a pathogenic
CC organism such as A. fumigatus, to treat a non-infectious disease in a
CC subject (e.g. cancer), to prevent or contain contamination of an object
CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
CC biofilm comprising A. fumigatus. The polynucleotides are useful for
CC expressing recombinant protein for characterisation, screening or
CC therapeutic use, as markers for host tissues in which the pathogenic
CC organisms invade or reside, for comparing with the DNA sequence of A.
CC fumigatus to identify duplicated genes or paralogues having the same or
CC similar biochemical activity and/or function, for comparing with DNA
CC sequences of other related or distant pathogenic organisms to identify
CC potential orthologous essential or virulence genes, for selecting and
CC making oligomers for attachment to a nucleic acid array for examination
CC of expression patterns, for raising anti-protein antibodies, as an
CC antigen to raise anti-DNA antibodies or to elicit another immune
CC response, and for identifying polynucleotides encoding the other protein
CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This sequence represents a protein of one of the essential genes
CC of Aspergillus fumigatus of the invention.
XX
SQ Sequence 869 AA;

Query Match 34.6%; Score 1345; DB 6; Length 869;
Best Local Similarity 35.9%; Pred. No. 2.7e-122;
Matches 317; Conservative 123; Mismatches 236; Indels 208; Gaps 26;

QY 11 VLCAKNLAKXDFER-----LPDPFAKIVVDGSGQCHSTDTVKT 49
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DQ 23 VIAADGLYKRDVFRKSVILAILLVRLTSGAFAGPDPFAVATVGE-QTHTSVIKKT 81
QY 50 LDPKWNQHYDLVVGKTDSTITISVWNHKKIHKQAGFLGCVRL----- 93
```

Db 82 LNPYNEMFDRVNEDESILAIQIFDQKF-KKQOGFLGVINRAIGVIDLQMGDGESL 140  
QY 94 ----SNAISRLK-----DTGYQLDCKLNPSDTAVRGQIVVSLQOT----- 131  
Db 141 PIRHSCDVSRLTFRLLLLETEMLTDLKSN--DNLVHVGKLIINLSNLSTNTPTNQANG 198  
QY 132 --RRIKGGVVDRCGLLEN-----EGTVYED--SGPG-----RPLSC 166  
Db 199 LHRSHVQSSTS-----SGLVPQVAPSSGHPAASGTAPVDPASPSLNPQVPSTTRPSST 254  
QY 167 FMEBPA-----PYTDSGAAAG--GNCREFE-----SP 193  
Db 255 AAPASAGAASVNSHGSRNLSFEDSQRLPAGWERREDNLGETYVDHNTTITWTRP 314  
QY 194 SQDRLQRLRN-----PDVR--GSLQTPQNRPHGOSP----- 226  
Db 315 SSNTNEHAQSQREANWOLERRAHSQRLPEDRTGANSPLPSSQAQHTPPAGGSANAV 374  
QY 227 -----ELPEGYEORTVQGVFLHTQTGVSTWHDPR-----IPDLNSV 266  
Db 375 SMATGATTAGTGLPPEQWEORTPEGRPYFVDHNTTITWDFRQOYRMGQANGT 434  
QY 267 N-----CDELGLPPEGWYRSTVSGRIYFVDHNNRTTQTFRLHIMHQCQLKEPS 319  
Db 435 NTTIQQPVSGLPGLPSGWMRLNTARVYFVDHNTTITWDDPR----- 479  
QY 320 QPLPLPSEGSLEDELPQAQRYERDLVKLVRLHLSLOQPAQCHIEVSRBEIESY 379  
Db 480 ----LPS--SL-DQGV--QYKDRFRKLIYFRSQPAL-RIMSGQCHVKVRRNIFEDSY 529  
QY 380 ROIMKRPDKLKKLMKFRCEGLDGGVAREWLYLLCHEMLNPPYGLFOYGTDMLYL 439  
Db 530 AEIMQASDLKKRLMKFDGEDGLDYGLSGEFFLLSHEMFPFCLFEYSAHDNYTL 599  
QY 440 QINPDSSINPHLSYFHFVRIMGLAVFHGYINGGTFVPPYKQLGKPIQLSLSVDP 499  
Db 590 QINPHSGVNPHELVYFPIGRVWGLAIFHRFLDSFFIGAFYKMKLKKVSLQDMGVDE 649  
QY 500 ELKSLVWLILNDITPVLDTFFCVHNAFGRILQHEIKPGRNVPVTEENKKEYVRLVYN 559  
Db 650 DLHRLNTWINDLEGVLELTFSDDEKFERRTIDLPKGRDIPVTEENKAYVRLVTE 709  
QY 560 WRFMEGIEAQLAKQGNELIPQHLKPFQDKELEIIGGLKIDLNDKSNTRLKHCV 619  
Db 710 WKIVKRVEEQNAFWSGNELIPADLVNVDERELELLIGIADIDVDDMKKHTDYRGY 769  
QY 620 ADSNIVRFWQAVETFDERRARLLQVGTSTRVPLQGRKALOGSTGAAPRLFTIHLID 679  
Db 770 ESDVIONFWKIVRSWDAEQSRLLQPTTGTSRIPVNGFKDLOGSD--GPRRFTIEK-S 825  
QY 680 ANTDLNKAHTCFNRIDIPPYSEYKELYKLLTAVEETCGFAVE 723  
Db 826 GDPALPKSHGTGNRLDLPYKSYETLEHKVSIATVEETLGGQE 869

## RESULT 13

AA30949

ID AA30949 standard; protein; 854 AA.

XX

AC AA30949;

XX

DT 21-OCT-1999 (first entry)

XX

DE Murine E3 ubiquitin protein ligase protein.

KW

E3 ubiquitin protein ligase; h-E3 UPL; antiinflammatory; antidiabetic;  
immunosuppressive; neuroprotective; cytoskeletal; antiarthritic; cardiac;  
immunomodulator; antiviral; treatment; screening; gene therapy; cancer;  
inflammation; autoimmune disease; neurological disease; apoptosis;  
endothelial cell; proliferation; differentiation; angiogenesis; leukemia;  
peripheral vascular disease; hematopoietic disorder; arthritis; cachexia;  
pulmonary disorder; diabetes; viral infection; murine.

```
Db 232 SSTGSLPTNTNTSTSEATSGLIPTISGGSGPRELNTVSOAPLPQGEORV----- 287
Qy 204 LRNPVVRGSL-----OTPNRPHGHQSPELPEGEORTTVQGVYFLHQTGTGVTWH 255
Db 288 ----DOHGRVYVDHVEKETTWDRE-----PLPGWERRVDMGRIYVDHFTTTTQW 338
Qy 256 DPRIP-----RDLE-----NSVNCDELGPLPFGWEVR 281
Db 339 RPTLESVRNVEQWLORLOQAMQFQNRFYGNQDLFATSKQNKFEFPLGPLPGWEKR 398
Qy 282 STVSGRIYFVDHNRRTQTDRLHHIMNHQOLKEPSOPLP----- 324
Db 399 TDSNGRVYFVNHTRITQWEDR-----SQQLNE--KPLPEGWEMRTVDGIFYVDH 450
Qy 325 -----PSEGSLEDELPQRYERDLVOKLVLR-----HELSQLQPOAGHCRIEVSE 372
Db 451 NRRATTYIDPRTGKSALDNGPQIAYVRDFKAKVQYERFWCQQLAMPQ-----HIKITVTRK 506
Qy 373 EPEESYRQIMKMRPKDLKRLMKRPGREGIDYGGVAREWYLLCHEMLNPIYGLFQYS 432
Db 507 TUFEDSQQIMSPQDLRRRLWVIFPGBEGDYGAVAREWFFLSHENVIMYCLFEYA 566
Qy 433 TONIYMLQINPOSSINPDHLSYFHFVGRIMGLAVFHGYINGFTVPFYKQLLGKPIQLS 492
Db 567 GKDNVCLQINPASYINPDHLKYFRFTRGFIAALFHGKFDITGFSLPFYKRLNKPVLK 626
Qy 493 DLESVDPELHKSILWLENDITPV-LDHTFCVHNHAFGRILHOLKPNGRNVPTVEENKK 551
Db 627 DLESIDPEFYNLSIMWKNNIBECGLEMYFSDVKELGEEKHDLKPNNGNTLVTEENKE 686
Qy 552 EYVRLVYVNRVGRIGBAQFALQKGNELIPOHLLKPPDKLELIGLDKIDLDNKS 611
Db 687 EYRMVAEWLSEGVZEQOAFEGNEILPOQLOYFDAKEVLLCQWQIDLDNWR 746
Qy 612 NTRLKHCVADSNIVRFWQAVETFDSEERARLLQFVTGSTRVPLQGFKALQGSTGAAGR 671
Db 747 HAIYRHYTRTSKQIMWFQFVKIEIDNEKRMRLQLFVTGTCRPLVGGFADLNGSN--GPQ 803
Qy 672 LFTIHLIDANTNLKPAHCENRDIIPPYESYEKLEKLLTAVEECGFAVE 723
Db 804 KPCIEKV-GKENWLPKSHTCFNRLDPPYKSYEQLEKLEKLLFAITEEGFGQE 854

RESULT 14
AAV30948
ID AAV30948 standard; protein; 852 AA.
XX
AC AAV30948;
XX
DT 21-OCT-1999 (first entry)
XX
DE Human E3 ubiquitin protein ligase protein.
XX
KW E3 ubiquitin protein ligase; h-E3 UPL; antiinflammatory; antidiabetic;
KW immunosuppressive; neuroprotective; cytostatic; antiarthritic; cardiant;
KW immunomodulator; antiviral; treatment; screening; gene therapy; cancer;
KW inflammation; autoimmune disease; neurological disease; apoptosis;
KW endothelial cell; proliferation; differentiation; angiogenesis; cachexia;
KW peripheral vascular disease; hematopoietic disorder; arthritis; leukemia;
KW pulmonary disorder; diabetes; viral infection; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 275..306
FT Domain /label= WWI
FT /note= "WW protein interaction domain I"
FT 307..340
FT Domain /label= WWII
FT /note= "WW protein interaction domain II"
FT 386..420
FT Domain /label= WWIII
```

```
FT Domain /note= "WW protein interaction domain III"
FT 427..460
FT /label= WWIV
FT /note= "WW protein interaction domain IV"
FN WO3940201-A1.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-GB000353.
XX
PR 05-FEB-1998; 98US-0073839P.
XX 30-APR-1998; 98US-00070060.
XX (ZENE ) ZENECA LTD.
XX
PI Hustad CM, Ghildyal N;
XX
XX WPI; 1999-508506/42.
XX N-PSDB; AAZ09235, AAY09236.
XX
PT New human proteolytic accessory enzyme and its modulators useful for
PT treating disease conditions like inflammation or autoimmune diseases.
XX
PS Claim 1; Fig 3; 95pp; English.
XX
CC This invention describes a novel human polynucleotide (I) which encodes a
CC E3 ubiquitin protein ligase, h-E3 UPL (II). The products of the invention
CC have antiinflammatory, immunosuppressive, neuroprotective, cytostatic,
CC antiarthritic, immunomodulator, antidiabetic, antiviral and cardiant
CC activity. The products of the invention can also be used for treating
CC patients with disorders mediated by the biological and/or pharmacological
CC activity of h-E3 UPL. The nucleic acid sequences encoding E3 UPL are used
CC in expression systems as assay for agonists and antagonists for the E3
CC UPL protein. The E3 UPL protein is used in screening assays to identify
CC blockers and antagonists. They are also used in gene therapy. Specific
CC modulation of biological and/or pharmacological activity of novel h-E3
CC UPL via administration of a modulator or heterologous expression, is used
CC for treating physiological conditions like inflammation, autoimmune
CC diseases, neurological disease, apoptosis, endothelial cell physiology
CC (e.g., proliferation, differentiation), peripheral vascular disease,
CC angioneuroma, cancer, hematopoietic disorders, arthritis, cachexia,
CC leukemia, pulmonary disorders, diabetes and viral infection. The
CC nucleotide sequences which encode h-E3 UPL may also be employed in
CC analysis to map chromosomal location e.g., screening for functional
CC association with disease markers. They are also used as screening tools
CC in the identification of appropriate human subjects and patients for
CC therapeutic clinical trials. The sequences can also be used to detect the
CC presence of the mRNA transcripts in a patient or to monitor the
CC modulation of transcripts during treatment. This sequence represents the
CC human E3 ubiquitin protein ligase protein described in the method of the
CC invention
XX
XX Sequence 852 AA;
XX
Query Match 33.5%; Score 1301; DB 2; Length 852;
Beat Local Similarity 34.7%; Pred. No. 5,7e-118;
Matches 308; Conservative 126; Mismatches 241; Indels 212; Gaps 26;
Qy 6 KIRITVLCAL-NLAKKOPFRUPDPFAKIVVDGSGQCHSTDTVKNTLDPKKNQHYDLYVGK 64
Db 9 QLOITVTSALKENKQWFG-PSPIVEVTD--QSKKTEKCNNTNPFKKQPLTIVTVP 65
Qy 65 TDSITISVNNHKTHKK--QGAGFLGCVRLLSNAISRLKDT-----GYQR-----L 109
Db 66 VSKLHFRVMSHQTLKSDVLLGTAALDIYETLKSNNMKLEEVVVTLLQGGDKREPTETIGDL 125
Qy 109 DLCKLNPSDITDAVRGQIVVSLQT-----RDRIGTGGG-----VVDG 144
Db 126 SIC-----LDGLQLESEVVNTGETTCSESASQNDGCSKSDTRVSTNGSDDPEDAGAGEN 181
Qy 145 RGL-----LENEGTVYEDSGPGRP-SCFMEEPAPYTDSTGAAGGNCRFVES----- 192
```

Db 182 RRVGNNFSLNNG--FKSRPRPP-----SRPPPTPRPAPVNGSPATSSDSGSST 234  
 QY 193 -----PSQDRLQAOQLRNPDP 208  
 Db 235 GSLPPTNTNTSEGATGLIPLTISGSGPRPLNPVTQAPLPFGWEQV-----D 286  
 QY 209 VRGSL-----QTPQNRPHGHSPELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP 260  
 Db 287 QHGRVYVVDHVEKRTWDRPE-----PLPFGWERRVDNMGRIYVVDHFTRTTTWQRTLE 341  
 QY 261 -----RDLE-----NSVNCDELGLPLPGWEVRSTVSG 286  
 Db 342 SVRVNEQWQLORSLOQAMQFNQRFYVGNQDLFATSSQKFEFPLGLPFGWEKRTSNG 401  
 QY 287 RIYFVDHNNRTQFTDPLRHIMHQCQLKEPSPQLP----- 324  
 Db 402 RYFVNNHTRITQWEDPR-----SQQLNE--RPLPEGWEMRFTVDGIPYFVDHNNRTT 453  
 QY 325 ----PSEGSLEDELPAAQRYERDLVOKLVLR---HELSQLQOQAHCRIEVSREEIFE 377  
 Db 454 TYIDERTKSALDNGPQIAYVRDPKAKVQYFRCQQLAMPQ-----HIKIVTKILFED 509  
 QY 378 SYRQIMKMRPKDKRLKRLMVKFRGEGLDYGVAREWLYLLCHEMLNPPYGLFOYSTDNII 437  
 Db 510 SPQIMSPSQDLRRRLRWIIPFGEGLDYGVAREWFLLSHEVLNPMYCLFEYAGKDN 569  
 QY 438 MLIQIPDSINPDHLSYFHFVGRIMGLAVFHGVINGFTVPVYKOLLGKPIQLSDLESV 497  
 Db 570 CLQINPASYINPDHUKYFRFGRITAMALFGKFDITGFSLFFRYRILNPKVGLKLESI 629  
 QY 498 DPELHKSILWILENDITPV-LDHTFCVNEHNAFRILQHELKPNGRNVPVTEENKKEYVRL 556  
 Db 630 DPEFNSLIWKENNIBECDELMYFSDVDEIKELSHDLKPNNGNINLVTEENKEEYIRM 689  
 QY 557 VYWRPFMGIAQFALQKGNELIPQHLKPFQKELELIIGLDKIDNDKWSNTRLK 616  
 Db 690 VAEMWLSRGVEQTOAFEGFNEILPQOYLQYFQAKSEVLVLLCGMIDNDWQWHAIR 749  
 QY 617 HCVADSNIVRWFQAVETFDERRARLIQVGTSTVPLQGFKALQSTGAAGPRLFTIH 676  
 Db 750 HYARTSKQIMWFQVFKVIDNEKRMLLQVGTGCRPLPVGGFADLMGNS--GPKFCIE 806  
 QY 677 LIDANTDLPKATCFNRIDIPPVESYEKLYEKLITAVEETCGFAVE 723  
 Db 807 KV-GKENWLPKRSHTCFNRLDLPYKSYEQKLEKLLFAIETEGRGQE 852  
 RESULT 15  
 ID ABR64210 standard; protein; 903 AA.  
 AC ABR64210;  
 DT 15-OCT-2003 (first entry)  
 DE Angiogenesis protein BNO104.  
 KW Cytostatic; antitumour; antiatheritic; antidiabetic; ophthalmological;  
 KW antipsoriatic; antiarteriosclerotic; cardiant; vasotropic; angiogenesis;  
 KW gene therapy; vasculature; cancer; rheumatoid arthritis; psoriasis;  
 KW diabetic retinopathy; cardiovascular disease; atherosclerosis;  
 KW ischemic limb disease; coronary artery disease.  
 OS Homo sapiens.  
 FN WO2003027285-A1.  
 PD 03-APR-2003.  
 PF 19-SEP-2002; 2002WO-AU001282.  
 PR 27-SEP-2001; 2001AU-00007973.  
 PR 27-SEP-2001; 2001AU-00007974.

PR 11-OCT-2001; 2001AU-00008210.  
 PR 29-OCT-2001; 2001AU-00008532.  
 PR 13-NOV-2001; 2001AU-00008838.  
 PR 28-AUG-2002; 2002AU-00951032.  
 XX (BION-) BIONOMICS LTD.  
 PA  
 XX Gamble JR, Hahn CN, Vadas MA;  
 FI WPI; 2003-354655/33.  
 XX DR N-PSDB; ACF34485.  
 XX  
 PT New angiogenic genes and polypeptides, useful for diagnosing,  
 PT prognosticating or treating an angiogenesis-related disorder, e.g.  
 PT cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or  
 PT cardiovascular diseases.  
 XX  
 PS Claim 15; SEQ ID NO 145; 90pp; English.  
 XX  
 CC The invention relates to the isolation of novel genes (ACF34446-ACF34559)  
 CC encoding proteins (ABR64180-ABR64281) involved in the process of  
 CC angiogenesis. The nucleic acid molecules are useful in identifying and/or  
 CC obtaining full-length human genes involved in an angiogenic process. The  
 CC nucleic acid molecule, polypeptides or complexes encoded, cells or  
 CC genetically modified non-human animals derived from these are useful for  
 CC the screening of candidate pharmaceutical compounds used in treating  
 CC angiogenesis-related disorders. They are also useful for diagnosing,  
 CC prognosticating or treating an angiogenesis-related disorder, which  
 CC involves uncontrolled or enhanced angiogenesis or is a disorder in which  
 CC a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis,  
 CC diabetic retinopathy, psoriasis or cardiovascular diseases such as  
 CC atherosclerosis), or involves inappropriately arrested or decreased  
 CC benefit (e.g. ischemic limb disease or coronary artery disease). The  
 CC modulator of expression or activity of the polypeptide encoded by the  
 CC nucleic acid sequence is useful for manufacturing a medicament for the  
 CC treatment of an angiogenesis-related disorder. This sequence corresponds  
 CC to one of the novel angiogenic protein  
 XX  
 SQ Sequence 903 AA;  
 Query Match 33.2%; Score 1288.5; DB 6; Length 903;  
 Best Local Similarity 33.8%; Pred. No. 1.1e-116;  
 Matches 312; Conservative 128; Mismatches 239; Indels 245; Gaps 29;  
 QY 6 KIRLTVLCAK-NAKAKDFRLLPDPFAKIVDGGSGCHSTDTVKNLTDPKWNQHYDLYGK 64  
 Db 19 QLQITVISAKLENKKNWFG-PSYVEVTVD--GQSKTEKCNNTNPKWQPLTVITVP 75  
 QY 65 TDSITISVNHKKIHK-----KQAGFLG-- 88  
 Db 76 VSKLHFRVWSHQTLKSDVLGTAALDIYETLKNWNKLEVVVVVLLQGLGDKPETTIGDL 135  
 QY 89 --CV-----RLLSNAISRLKDTGYQ---RLDL-----CKL-----NPSDTPAV 121  
 Db 136 SICLDGLQLESEVINGETTCSENGVSLCLPFLCNSAISAHCNLCPLGSDSPISASRV 195  
 QY 122 RGQIVVS-----LQTRD--RIGTGS-----VVDGRGL-----LENEGTVVEDSGP 160  
 Db 196 AGFTGASQNDGDSRSDKDETRVSTNGSDPDADAGENRRVSGNNSPFLSNGG--FKPSRP 253  
 QY 161 GRPLSCFMEEPAPYTDSTGAAAGGNCRFVES----- 192  
 Db 254 PRP-----SRPPPTPRPAPVNGSPATSSDSGSSTGSLPPTNTNTSEGATSLIIP 308  
 QY 193 -----PSQDRLQAOQLRNPDPVGRSL-----QTPQNRPHG 223  
 Db 309 LTAISGSGPRPLNPVTQAPLPFGWEQV-----DQGRVYVVDHVEKRTWDRPE-- 358  
 QY 224 QSPPELPEGYEORTTVQGVYFLHTQTGVSTWHDPRIP----- 260  
 Db 359 ----PLPFGWERRVDNMGRIYVVDHFTRTTTWQRTLESVRVNEQWQLORSLOQAMQFN 415

```

261 QY -----RDL-----NSVNCDELGLPLPGQWEVRSTVSORIIYFVDHNRTRTQFTDPLRHHM 309
416 DB QRIYVGNQDLFATOSKBEFDPLGPLPGQWEKRTDSNGRVFVNHNTRITQWDEPR-----470
310 QY NHQOQLKEPSQPLPL-----PSEGSLEDEELPAQRYBERD 343
471 DB -SQQLNE--KPLPSGWMRFTVDGIPYFVDHNRRTTITDPTGKSALDNGPQIAYVRD 527
344 QY LVQKLKVLK--HELSQLQOPQAGHCRITEVSEETFEESYRQIMKMPKOLKXELMWKRG 400
528 DB FRKVQYFRWCQQLAMPO---HIKTIVTKTLFEDSPQOIMSFQPDRLRRLWIFPG 583
401 QY EEBGLDYGVARWLYLICHEMLPNPYIGLFQYSTDNIIYMLQINPDSSINPDHLSYHFYVGR 460
584 DB EEBGLDYGVARWFFLLSHEVLNPMWCLFYVAGKDNVCLQINPASYINPDHLKYFEYGR 643
461 QY IMGLAVFHGHYINGFTVPFYKQLLGRPIQLSDLESVDPELHKSLLWILIENDITPV-LDH 519
644 DB FIAMALFHGKFIDTGFSLPFYKIRLKNLKVGLKJDESIDPEFYNSLWVXKNNIECEDLEB 703
520 QY TFCVEHNARFGRILQELHQLPKNRNVPTBENKKEYRLYVNVNRMFRMGIEAQFALAQGFNE 579
704 DB YFSVDKEIILGBIKSHDLKPNNGNIIIVTEENKEEYIRMVAEWRLSRGVEBOTQAFEGFNE 763
580 QY LIPQHLKFPDOKELIILIGLDKIDINDWKSNTRLKXHVADSNIIYVWFWQAVETDEBR 639
764 DB ILPQQYLOVFDAKELEVLGCGMQEIDLNDWQORHAIYRYHARTSKQIMFWFQFYKEDNEK 823
640 QY RARLLQFVTGSRVPLQGFKALQOGSTGAAGPRLFTIHLIDANTDNIIPKHAHTCFNRIDIPP 699
824 DB RMELLOFVTGTCELPUVGFADLMSGN---GPOKFCIEKV-GKENWLPFRSHTCENRLDLP 879
700 QY YSSEYKLYEKLLTAVEETCGFAVE 723
880 DB YKSVEQLKEKLLFAIETEFGFQOE 903

```

Search completed: September 21, 2004, 07:43:38  
Job time : 99.4534 secs

This Page Blank (uspio)

64 NIDSIIISVWNHAKIHKARQGAGFTGCVRLTSLNALSRLKDTGYQRL--DLCKLNPSDTDAV 121

Db 66 DNSTIAIQVFDOKKF-KKKGGQFLGVINLRVGVLDLALAGDEMLTRDLKKN--ENTVV 122  
 QY 122 RGQIVVSLQ-----TRDRIGTGGSVVDCRGLLENEGTVEY 157  
 Db 123 HGKIIINLSTTAQSTLQVPSAASGARTQRTSITNDPOSSQSSSVSRNPASSRAGSPTRD 182  
 QY 158 SGPG-----RPLSCFME-----PAPY-----TDSTG-----AAA 182  
 Db 183 NAPAASPASSERTFSSFDQGRPLPGWERTDNLGRYVVDHNTSTRITWIRNLSVA 242  
 QY 183 GGCNCRFVSPS-----QDRLQARLNPDPVRSGLQTPNRPNGHQSPPELPGYEQ 234  
 Db 243 GAAAELHSSASSANVTGVQPSSSNAARTEASVLTSNATTAG-----SGELPPGWEQ 296  
 QY 235 RTTVQGVYVLTQTCVSTWHDPRI PRDLNSV-----CDELGPLPGWEVRS 283  
 Db 297 RYTPGGRPYVDHNTKTTTWDPRQYIRSYGGNNATTIQQQVPSQGLPSPGEMRLT 356  
 QY 284 VSGRIYFVDHNRITQFTDPRLHHIMNHQCOLKEPSQPLPSPSGSLEDELPQAQRYERD 343  
 Db 357 NTARVYFVDHNTKTTTWDPR-----LPS--SL--DQNV--QYKRD 392  
 QY 344 LVOKLVLRHLSLQOQAGCRIEVSREEIFESYRQIMKRPKDLKRLMKVFRGEEG 403  
 Db 393 FRKLIYFSLQAL-HPLPGQCHIKVRRNHIFEDSYABIMRQSATDLDKRLMKVFRGEEG 451  
 QY 404 LDYGGVAREWLYLICHEMLNPYYGLFOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMG 463  
 Db 452 LDYGLSREYFLLSHEMFPYCLFEYSVDNYTLQINPHSGINPEHLNFKFGRVIG 511  
 QY 464 LAVFHGYINGGTFVPFYKOLLGKPIQLSDLESVDPELHKSILWILENDITPVLDHTFCV 523  
 Db 512 LAIFHRFRVDAFFVSVFYKMLQKVTLDQMSDAEYVSLWILNDITGVLDLTFV 571  
 QY 524 EHNAFGRILQHELKPNRNPVTBENKEEYVRLVYNNWFRMGIEBAQFLALOKGNEFLIPQ 583  
 Db 572 EDNCFGEVVTIDLPKNGRNIETVEENKEEYVDLVTVWRIQRIEEOFNAPHEGSELLIPQ 631  
 QY 584 HLLKPPDOKELELIGGLDKIDLNDKSNTRKLVKCHVADSNIVRWFWQAVETFDERRARL 643  
 Db 632 ELINVFDERELELLIGGSEIDMEDWKKHDKYRSYSENDQIWKFWELMDWSNEKKSRL 691  
 QY 644 LQFTVGTSTVPLQGFALQGSTGAAGPRLFTIHLIDANTNLPAKTCFNRDIPPPYESY 703  
 Db 692 LQFTTGTSRIPVNGFKOLOGSD--GPRKFTIEKA-GEPNKLPKAHTCFNRDLDPPTYSK 747  
 QY 704 EKLYEKLITAVEETCGFAVE 723  
 Db 748 KDLDHKLSTIAVEETIGFGQE 767

RESULT 2  
 S66562  
 ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)  
 N:Alternate names: E6-AP-like protein ubiquitin ligase  
 C:Species: Schizosaccharomyces pombe  
 C:date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002  
 C:Accession: S66562; T45159  
 R:Nefsky, B.; Beach, D.  
 EMBO J. 15, 1301-1312, 1996  
 A:title: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25  
 A:Reference number: S66562; MUID:96205868; PMID:8635463  
 A:Accession: S66562  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-766 <NEF>  
 A:Cross-references: GB:U66716; NID:g1519443; PIDN:AAB07514.1; PID:g1519444  
 R:Nefsky, B.S.; Beach, D.  
 submitted to the EMBL Data Library, August 1996  
 A:Description: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of  
 A:Reference number: 222935  
 A:Accession: T45159  
 A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA  
 A:Residues: 1-766 <NE2>  
 A:Cross-references: EMBL:Y07592; PIDN:CAA68867.1  
 C:Genetics:  
 A:Gene: publ  
 C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
 C:Keywords: ligase  
 F:205-242/Domain: WW repeat homology <WW1>  
 F:288-325/Domain: WW repeat homology <WW2>  
 F:345-382/Domain: WW repeat homology <WW3>  
 Query Match 37.9%; Score 1471; DB 1; Length 766;  
 Best Local Similarity 40.9%; Pred. No. 6.7e-99;  
 Matches 327; Conservative 118; Mismatches 235; Indels 120; Gaps 20;

QY 4 SIKIRLTVLCAKNLAKDPFRLLPDPFAKIVVDSSGQCHSTDTVKNVLDPKWNOHYLDLYVG 63  
 Db 7 SRRIRVTIAADGLYKRDVFRFPDPFAVLTVDDGE-QTHTTTAKKTLNPNYNNETFEVNT 65  
 QY 64 KTDSTISVWNNHKKIHKKGAGFLGCVRLLSNAISRLLKDTGYQRL--DLCKLNPSDITAV 121  
 Db 66 DNSTIAIQVFDOKKF-KKKGGQFLGVINLRVGVLDLALAGDEMLTRDLKKN--ENTVV 122  
 QY 122 RGQIVVSLQ-----TRDRIGTGGSVVDCRGLLENEGTVEY 157  
 Db 123 HGKIIINLSTTAQSTLQVPSAASGARTQRTSITNDPOSSQSSSVSRNPASSRAGSPTRD 182  
 QY 158 SGPG-----RPLSCFME-----PAPY-----TDSTG-----AAA 182  
 Db 183 NAPAASPASSERTFSSFDQGRPLPGWERTDNLGRYVVDHNTSTRITWIRNLSVA 242  
 QY 183 GGCNCRFVSPS-----QDRLQARLNPDPVRSGLQTPNRPNGHQSPPELPGYEQ 234  
 Db 243 GAAAELHSSASSANVTGVQPSSSNAARTEASVLTSNATTAG-----SGELPPGWEQ 296  
 QY 235 RTTVQGVYVLTQTCVSTWHDPRI PRDLNSV-----CDELGPLPGWEVRS 283  
 Db 297 RYTPGGRPYVDHNTKTTTWDPRQYIRSYGGNNATTIQQQVPSQGLPSPGEMRLT 356  
 QY 284 VSGRIYFVDHNRITQFTDPRLHHIMNHQCOLKEPSQPLPSPSGSLEDELPQAQRYERD 343  
 Db 357 NTARVYFVDHNTKTTTWDPR-----LPS--SL--DQNV--QYKRD 392  
 QY 344 LVOKLVLRHLSLQOQAGCRIEVSREEIFESYRQIMKRPKDLKRLMKVFRGEEG 403  
 Db 393 FRKLIYFSLQAL-HPLPGQCHIKVRRNHIFEDSYABIMRQSATDLDKRLMKVFRGEEG 451  
 QY 404 LDYGGVAREWLYLICHEMLNPYYGLFOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMG 463  
 Db 452 LDYGLSREYFLLSHEMFPYCLFEYSVDNYTLQINPHSGINPEHLNFKFGRVIG 511  
 QY 464 LAVFHGYINGGTFVPFYKOLLGKPIQLSDLESVDPELHKSILWILENDITPVLDHTFCV 523  
 Db 512 LAIFHRFRVDAFFVSVFYKMLQKVTLDQMSDAEYVSLWILNDITGVLDLTFV 571  
 QY 524 EHNAFGRILQHELKPNRNPVTBENKEEYVRLVYNNWFRMGIEBAQFLALOKGNEFLIPQ 583  
 Db 572 EDNCFGEVVTIDLPKNGRNIETVEENKEEYVDLVTVWRIQRIEEOFNAPHEGSELLIPQ 630  
 QY 584 HLLKPPDOKELELIGGLDKIDLNDKSNTRKLVKCHVADSNIVRWFWQAVETFDERRARL 643  
 Db 631 ELINVFDERELELLIGGSEIDMEDWKKHDKYRSYSENDQIWKFWELMDWSNEKKSRL 690  
 QY 644 LQFTVGTSTVPLQGFALQGSTGAAGPRLFTIHLIDANTNLPAKTCFNRDIPPPYESY 703  
 Db 691 LQFTTGTSRIPVNGFKOLOGSD--GPRKFTIEKA-GEPNKLPKAHTCFNRDLDPPTYSK 746  
 QY 704 EKLYEKLITAVEETCGFAVE 723  
 Db 747 KDLDHKLSTIAVEETIGFGQE 766

RESULT 3

[illegible][illegible]

db 174 RNHSTSHPSRGTAQAVESTIOSGTTAATNTATTSHPSNSTSSATPOVSSPEDOVCPIBP 223

```
QY 183 -----GGGCRNCFVE-----SPSQDRLQAQRLRNPDVRSGLQTPQNR-----PHG 222
Db 234 GWERRDNFGRTYVVDHNTTRTTWKPTLDQ--TEAERGQNLNANTELERRQHRGRTLPGG 292
QY 223 HQS-----PELPGYEQRTTVGQVYFLHT 247
Db 293 SSNSSVTVQGGSNIPVNGAAAAFAATGTTTSGLGELPSGWQRTTPGRTAYVDH 352
QY 248 QTGVSTWHDPRIPRLDLSVN-----CDELGLPFGWEVRSTVSGRIYFVDHNNRT 297
Db 353 NRTTITWDRRQYHRTYGTNTTQQQPVSQLGLPSGWEMLNTARVYFVDHNTKT 412
QY 298 TQFTDRLHMHQCKLPSPQLPLPSGSLDEELPAQRYERDVLQKLVLRHLSL 357
Db 413 TTWDDPR-----LPS--SL-DQNV--QYKDRFRKVIYFRSQPAL 448
QY 358 QPQAGHCRTEVSGREEIFEESYRQIMQRPKDLKRLMKVFRGEEGLDYGGVAREWLYLL 417
Db 449 -RLIPGOCHIKVRKNTIFDAYQEIWRQTPEDLKRLMIKFDGEEGLDYGGVSRFFELL 507
QY 418 CHEMLNYYGLFOYSTDNIMYLNPDSSINPDHLSYFHFVGRIMGLAVPHGYINGGT 477
Db 508 SHEMFNPFYCLFEYSAYDNTYTIQINPNSGINPBLAYFKFIRGVWGLGVFHRFLDAFV 567
QY 478 VPFYKQLLGKPIQLSDLESVDPELHKSLVILENDITPVLDTHTFCVEHNAFGRILQHELK 537
Db 568 GALYKMLRKKVVLQDMEGVDAEVYNSLWMLNLESDGVLDTLFSADDERFGEVTVDLK 627
QY 538 PNGRNVPVTBENKKEYVRLVYNNFRMRGIEBAQFLALQKGNELIPQHLKPPQKLELI 597
Db 628 PDGRIEIVTGNKKEYVELYVQWRIVDVQVQKAFMDGNGNELIPEDLVTVFDERELELL 687
QY 598 IGGDLKIDLNDKSNTRLKHCVDNSNTRVWFMQAVETFDERRARLLQFVTGSTRVPLQG 657
Db 688 IGGFAEIDIEDWKHTDYRGQESDEVIOFWKCKVSEWNEQEARLLQFTTGTSTRIPVNG 747
QY 658 FKALQGSTGAAGPLFTIHLIDANTDNLPAKHTCFNRIDIPPEYSEYKLYEKLTLAVEET 717
Db 748 FKDLQGSQD---GPRRFTIERA-GEVQQLPKSHTCFNRVDLPQVVDYDSMKQKLTAVEET 803
QY 718 CGPAVE 723
Db 804 IGFQGE 809
RESULT 5
T39585
ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T39585
R:Volckaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21865
A:Accession: T39585
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-786 <VOL>
A:Cross-references: EMBL:Z99759; PIDN:CAB16903.1; GSPDB:GNO0067; SPDB:SPBC16E9.11c
A:Experimental source: Strain 972h; cosmid c16E9
C:Genetics:
A:Gene: SPDB:SPBC16E9.11c
A:Map position: 2
A:Introns: 60/2
C:Superfamily: Yeast ubiquitin-protein ligase; WW repeat homology
F:236-273/Domain: WW repeat homology <WW1>
F:306-343/Domain: WW repeat homology <WW2>
F:364-401/Domain: WW repeat homology <WW3>
Query Match 34.5%; Score 1341; DB 2; Length 786;
Best Local Similarity 36.6%; Pred. No. 2.1e-89;
Matches 304; Conservative 128; Mismatches 236; Indels 162; Gaps 21;
```

```
QY 6 KIRLTVLCAKNLAKKOFRLPDPAKIVVDGSGOCHSTDTVKNTLDPKMKHQHYDVLVGKT 65
Db 7 KRVRYIVAAOGLSKRDURFPQDFPAILTVDGE-QTHYTKVKKSVAPYWNSEGFEVTVKPS 65
QY 66 DSITISVMNHKKHKKGAGFLGCVRLLSNAISRLKDTGYQR-----LDLCK-LNPSDTD- 119
Db 66 SVISIRLFDQKF-KKXQOGLGLVSP-----RMREVGSGFRSNREVSLTEPLKKKSTTN 118
QY 120 -AVRGOIIV-----SL 129
Db 119 LSVLGNLVLKVAPSKIRAPAGNHSSTTANRTTPTTTTARTTTPRPTATNTNQST 178
QY 130 QTRDRIGTGSVVDRCGLLENEGTVYEDSGPGRPL-----SCFMEEPA-----PYTDS 177
Db 179 SNSFTNGTSAATNGTGTGAGTGASHSS-----PVTNRQTNNTSALNSNAHIMSSPEDQ 234
QY 178 TGAAGG-----GNCRFVE-----SPSQDRLQAQRLRNPDV 209
Db 235 YGRLLPPGWERRADSLGRITYYVDHNTTRTTWTRPASSNPNVHNTSSDSQRLNHQNRHLPD- 293
QY 210 RGSLLQTPQNRPHGHOSP---ELPEGYEQRTTVGQVYFLHTQTGVSTWHDPRIP- 260
Db 294 -----DSNPSLQMSQSDGNDLPGWENRYTDTGRPYFVDHNTTITTTWVDPNPLVRPNG 346
QY 261 -----RDLNSVNCDELGLPPLPGWENRSTVSGRIYFVDHNNRTTOFTDPRLHHMHQCCQ 314
Db 347 GSSTVGSMLQPQSLSHLGLPLSGWEMBLTNSARYFVDHNTKTTTWDPR- 397
QY 315 LKESQPLPLPSEGLSEDEELPAQRYERDLVQKLVLRHLSLQOPOAGHCRIEVSRREI 374
Db 398 -----PSALDQDVP--QYKCDFRKLIYFRSQPM-RPLPGQCNVVRDHI 441
QY 375 FEESYRQIMKMRPKDLKRLMKVFRGEEGLDYGGVAREWLYLLCHEMLNPPYGFQISTD 434
Db 442 FEDSYAEIMRYSAHDLKRLMIRFDGEDGLDYGLSREFFFLSHKMFDPITYCLFEYSAV 501
QY 435 NIYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHGHVINGGTVPFYKOLLGKPIQLSDL 494
Db 502 DNYTQINPHSINPEHLNYPFRFGRVIGLAIHRRFLDAFFVLSLYKLLRKKVSLADM 561
QY 495 ESDVDELKHSLVVILENDITPVLDTHTFCVEHNAFGRILQHELKPNGRNPNVTEENKKEYV 554
Db 562 ESIDAEFYRSLKWLLENDITGLDITFSVEEDHGEVTRVELITNGENIEVTEENKKEYV 621
QY 555 RLYVNRWFRMGIEAQLQKGNELIPQHLKPFQDKLELITGLDKIDLNDKSNTR 614
Db 622 DLVTEWRYSKRVEQQFNAYFGFVELVSPDLVNVFDERELLELLGGISDVEDVDWKSHTE 681
QY 615 LKHCVDNSIVWFMQAVETFDERRARLLQFVTGSTRVPLQGFKALQGSAGAPRLFT 674
Db 682 YRTYATDPVVKWFWEIAGWKQEDRSKLLQFAIGTSIRIPVNGFEDLQGSQD---GPRKFT 738
QY 675 IHLIDANT-DNLPAKHTCFNRIDIPPEYSEYKLYEKLTLAVEETCGPAVE 723
Db 739 IE--KAGTPDQLPVAHTCFNRDLDPDPSKDTLHEKLSLAVENTVGFQNE 786
RESULT 6
T46412
ubiquitin-protein ligase (EC 6.3.2.19) NEBD4 - human (fragment)
N:Alternate names: hypothetical protein DKFZp434P2422.1
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 03-Jun-2002
C:Accession: T46412
R:Blum, H.; Bauseachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23034
A:Accession: T46412
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-820 <AAA>
A:Cross-references: EMBL:AL137469
```

A:Experimental source: adult testis; clone DKFZp434P2422

C:Genetics:

A:Gene: GDB:NEDD4

A:Note: DKFZp434P2422.1

C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin

C:Keywords: ligase

F:142-179/Domain: WW repeat homology <WWR1>

F:342-379/Domain: WW repeat homology <WWR2>

F:393-430/Domain: WW repeat homology <WWR3>

F:489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 32.0%; Score 1241.5; DB 2; Length 820;

Best Local Similarity 34.9%; Pred. No. 3.9e-82;

Matches 294; Conservative 116; Mismatches 230; Indels 203; Gaps 21;

QY 43 TDTVKNTLDPKKNQHYDLYVGKTD--SITISVNNKHKHKQAGFLGCVRLLSNAI----- 97

DB 12 TKTIKTLNPKNBEFYFVNPSNRLLFEVDENLRTRDD--FLGQVDVPLSLHPTED 68

QY 98 -----SLKDTGYQRLDLCLNPSDDTDVARGQIIVSLQTRDRIQT 137

DB 69 PTWERPYTFKDFLLPRSHKSVK--GFLRLKWAYPKNG-----GQDEENSQDRDMEH 121

QY 138 GGSVVDRCG-----LLENEGTVY----- 155

DB 122 GWEVDNDSASQHQEELPPPLPGWEEKVDNLRGTYYVNNHNRRTQWHRPSLMDVSE 181

QY 156 -----EDSGP-----GRPLSCFMEEPAPYTDSTGAA-- 181

DB 182 SDNNIRQINQEAARRFRSRHHISEDLPEPSEGQDVPEPWETISEEVNIAGDSLGLALP 241

QY 182 ---AGGNCRFVESPSQD-----ORLQAQLRN----- 206

DB 242 PPPASFGSRTSQELSELSRLRLQITPSNGSQFSLIQREPSSLRSCSVTDVAEQGH 301

QY 207 -----PDVRGSLQTPNRRPHGHOSPE-----LPGEYEQRTTVQGVFL 245

DB 302 LPFGAKDSPVRAVKDTLSNPSPQSPSPYNSPKPHKVTQSLPFCWEMRIAPNGRPF 361

QY 246 HTQTGVSTWHDPRIP-----RDLNVNCDLGLPFGWEVSTVSGRIYFVDHNNRTTQF 300

DB 362 DHNTKTTTWEDPRLKFPVPMHRSKTSLNPNLGLPFGWEERHLDGRTFYIDHNSKIQTQ 421

QY 301 TDPRLHHIMHQCQLKEPSQLPLSEGSLEDEELPAQRYERDLVKLVLRHLSLOQP 360

DB 422 EDFRLQN-----PAITG-----PAVPYGRFKQKYDYFRKKLKKPAD 458

QY 361 QAGHCRIEVSRBEEIPEESYRQIMKM-RPKDKRLMVKFRGBEGLDYGGVAREWLYLLCH 419

DB 459 IPNRPEMKLHNNIIEESYRIRMSVKRPDVLKARLWIEFESEKGLDYGVAREWFLLSK 518

QY 420 EMLNPPYGLFQYSTNIMYLNQINPDSSI-NPDHLSYFHFVGRIMGLAVFHGHIYNGGTV 478

DB 519 EMFNPPYGLFEYSATDNTYLTQINPNSGLCNEPDHLSYFTFICRVAGLAVFHGKLDDGFFIR 578

QY 479 PFYKOLLGKPIQLSDLESVDPLHKSLSVILENDITPVLDHTEFCVEHNAFGRILQHELKP 538

DB 579 PFYKMWLGKQITLNDMESVDSEYVNSLKWILENDPTE-LDLMFCIDENFQCTQVQDLKP 637

QY 539 NGRNVPTTEENKKEYVRLVYNNRFRGIEAQFLAQKGFNELIPOHLLKPFQDKELELI 598

DB 638 NGEIMVTNENREVIDLVIQWRVNRVQKQMNALFEGTELLPIDLIKIFEDNELELLM 697

QY 599 GGLDKIDLNDKSNRLKHKVADSN-IVRWFWQAVETEDERRARLLQFVTGSTRVPLOQ 657

DB 698 CGLGDVDNDWQHSYKNGYCNPFVPIQWFQKAVLLMDASKIRLLQFVTGSTRVPMNG 757

QY 658 FXALOGSTGAAPRLFTIHLIDANTDNLKFAHTCFNRIDIPPESEYKLYEKLTAVEET 717

DB 758 FAELYGSN--GPQLFTIEQW-GSEPKLPRAHTCFNRDLDPPEYTFEDLREKLLMAVENA 813

QY 718 CGF 720

DB 814 QGF 816

#### RESULT 7

S70642

ubiquitin ligase Nedd4 - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 18-Aug-2000

C:Accession: S70642

R:Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.

EMBO J. 15, 2371-2380, 1996

A:Title: WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na(+)

A:Reference number: S70642; MUID:96221297; PMID:8665844

A:Accession: S70642

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-887 <STA>

A:Cross-references: EMBL:U50842; NID:G1293646; PIDN:AA84949.1; PID:G1293647

C:Genetics:

A:Gene: Nedd4

C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin

F:54-167/Domain: proline kinase C C2 region homology <KC2>

F:245-283/Domain: WW repeat homology <WWR1>

F:402-439/Domain: WW repeat homology <WWR2>

F:459-496/Domain: WW repeat homology <WWR3>

F:555-881/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 31.6%; Score 1227.5; DB 2; Length 887;

Best Local Similarity 33.9%; Pred. No. 4.6e-81;

Matches 296; Conservative 127; Mismatches 224; Indels 227; Gaps 28;

QY 7 IRUTVLCANLAKKDPFRLEPPFAKIV---VDGSGQCHSTDTVKNLTDPKN----- 55

DB 77 VRVKVIAGIGLAKKILGASDPYRVTVLDPMSGVLTSVQTKIKKSLNPKNWEILFRV 136

QY 56 ---QH-----YDLYVGKTD--SITISVMN-----HKHKHKKQ 83

DB 137 LPQHRILFEVDENLRTRDDFLGQVDVPLYPFTENPRMERPYTFKDFVLHPRSHKSRV 196

QY 84 AGFLGCVRLLSNAISRLKDT-----RLKMTYLPKNGSDDENADQAELEPGWVVLQDPDAATHLQHPPEP 245

DB 197 KGYL-----KLNPDL-----KLNPDL-----KLNPDL-----KLNPDL----- 103

QY 104 -----GY-QRLDLC-----KLNPDL-----KLNPDL-----KLNPDL-----KLNPDL----- 135

DB 246 SPLPFGWEERQDVLGRYVYVNHESRTTQMKRPSDEDDLTDDENGDIQLQAHGAFTTRQI 305

QY 136 GTGGSVVDRCG-----LLENEGTVYEDSGPGRPLS-----C 166

DB 306 SE-----DVGDPDNHSPENWEIVREDNTIYSGAVQSPSPGHDPVQVRLABELTRLT 360

QY 167 FMEEPA---PYTDSFGAAAGGN--CRFVESPSQDQRLQALRNPDVGRSLQTPQNRPH 221

DB 361 MYGNPATSQPVTSNHSRSGSGTCTFEEQPTLPVLLPT----- 400

QY 222 GHOSPELPEGEYEQRTTVQGVYFLHTQTGVSTWHDPE-----RIPRDLNS---VNCDE 270

DB 401 ---SSGLPFGWEERQDGRSYVVDHNSKTTTWSKTMQDDPRSKIPAHLRGKTPTVDSND 457

QY 271 LGPLPFGWEVRSVSGRIYFVDHNNRTTQFTDPRLHHIMHQCQLKEPSQQLPLPSEGSL 330

DB 458 LGPLPFGWEERTHTDGRVFPFINIKKTQWEDPRMQNV-----AITGPAEP----- 503

QY 331 EDEELPAQRYERDLVKLVLRHLSLQDPOAGHCIEVSRBEEIPEESYRQIMKRPKD- 389

DB 504 -----YSRDKRYKRYEFREKRLKQKQTDIPNKFEMKLRANILSDSYRRINGVRADF 554

QY 390 LKRLMKVFRGEGLDYGGVAREWLYLLCHEMLNPPYGLFQYS-TDNIMYLNQINPDSSI- 447

DB 555 LKARLWIEFDGEGKLDYGGVAREWFFLSKEMENPYVGLFEYSATEDNTYLTQINPNSGLC 614

QY 448 NPDHLSYFHFVGRIMGLAVFHGHIYNGGTVFPYKOLLGKPIQLSDLESVDPLHKSLSVW 507

```
Db 615 NEDHLSYFKPIGRVAGMAVYHGLDGLDFFIRPPYKQMLQKLITLHDMESVDSEYYSLSLW 674
QY 508 ILNDITPVLDTFCVEHNAFGRILQHELKFNKRNVPVTEENKEVYVLYVNRWFRGIE 567
Db 675 ILNDPTE-JDLRFIDEELFGTHQHELKGTGSEVVVTKNKKKEIYLVIOQRVNRIO 733
QY 568 AQFLALQKGFNELIPQHLKPPQKLELIGGLDKIDLNDKSNRKLKHCVA-DSNIVR 626
Db 734 KQMAAFKEGFFELIPQDLIKIFDENELMLCGLGDVNDWREHTKYNGYSLNHQVTH 793
QY 627 WFWQAVETFEERRARILQVTCSTRVPLOGFKALOGSTGACGPRFLTILHIDANTDNL 686
Db 794 WFWKAVLMDSEKRIQLQVTCSTRVPMGFAELYGSN---GPGSFTVEQW-GTPDKLP 849
QY 687 KAHTCFNRIDIPPYSEYKELYEKLTAVENTCGF 720
Db 850 RAHTCFNELDLPYSEDFELWDLQMAIENTQGF 883

RESULT 8
NEDD-4 ORF - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 18-Aug-2000
C:Accession: 183196
R:Kumar, S.; Tomooka, Y.; Noda, M.
Biochem. Biophys. Res. Commun. 185, 1155-1161, 1992
A:Title: Identification of a set of genes with developmentally down-regulated expression
A:Reference number: I60167; MUID:92328780; PMID:1378265
A:Accession: 183196
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-708 <RES>
A:CROSS-references: GB:D10714; NID:g220508; PID:g220509
C:Genetics:
A:Gene: NEDD-4
C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin
F:40-77/Domain: WW repeat homology <WW1>
F:196-233/Domain: WW repeat homology <WW2>
F:251-288/Domain: WW repeat homology <WW3>
F:347-682/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 27.1%; Score 1051.5; DB 2; Length 708;
Best Local Similarity 36.1%; Pred. No. 2.2e-68;
Matches 244; Conservative 102; Mismatches 194; Indels 135; Gaps 19;

QY 50 LDPKWNQHYDLYVGKTSITISVWNKKHKKQAGLGCVRLLSNALSRLKDTGYQLD 109
Db 42 LPPGWEERQDV-LGRITYV---NHES-----RRTQWR-- 70

QY 110 LCKLNP-----DIDAVRGQIWSLQTRDRTGTGGSVDCRGLLEN-----EGTV 154
Db 71 -----PSFDDDLTDNDMDQLQRAFTTRQISEVDGPDNRESPEWNIIVREDENTE 125

QY 155 YEDSGGCRPLSCFME-----EPA---PYTDTGAAAGG--NCRPVES 192
Db 126 YSGQAVQSPSGHIDVQTHLAEENFTRLAVCGNPATSPQVTSNHSRGGSLQTCIFEEQ 185

QY 193 PSQORLQAOQLRNPVDRGSLQTPQNRPHGHSQELPEGEYQRTTQGVYVFLHTGTGS 252
Db 186 PTLPVLLPT-----SSGLPPGWEKQDDRGRRYYVDHNSKTT 222

QY 253 TWHDP-----RIPRDL-NSVNCDELGPLPGWEVRSTVSGRIYFVDHNNRTTQDTP 303
Db 223 TWSKETMQDDPSRKIPAHRLRGKTSNDLGPLPGWEERTHTDGRVFTINHNKKTQWEDP 282

QY 304 RLHIMNHQCLKEPSQPLPLPSGSLDEBELPAQRVERDLVQKLKVLRLHLSLQOQPAQ 363
Db 283 RLQNY-----AITGPAV-----YSRDYKRYEYFRRLKKKQTDIPN 319

QY 364 HCRTEVSREEIFESYROIKMKRPD-LKKELVMKFRGEBGLDYGGVAREWLVLICHEML 422
Db 320 KFEKRLRANILEDYRIRMGVKADLLKASLWIEFDEKGLDYGGVAREWFFLISKEMF 379
```

```
QY 423 NPYYGFLFOYSTDNYMLQINPDSSI-NPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFY 481
Db 380 NPYYGFLFEYSATDNYTLQINPNSGLCNEDHLSYFKPIGRVAGMAVYHGLDGLDFFIRPFY 439
QY 482 KQLLGKPTQLSDLESVDPELHKSILVLENDITPVLQHTFCVEHNAFGRILQHELKPNGR 541
Db 440 KQMLQKLITLHDMESVDSEYYSLSLWLENDPTE-JDLRFIDEELFGTHQHELKGTGGS 498
QY 542 NVPYTEENKKEYVLYVNRWFRMGIEAQFLALQKGFNELIPQHLKPPDQKLELELIIGL 601
Db 499 EIVVTNKNKKEIYLVIOQRVNRIOKMAAFKEGFFELIPQDLIKIFDENELMLCGL 558
QY 602 DKIDLNDKSNRKLKHCVA-DSNIVRWFQWQAVETFEERRARILQVTCSTRVPLOGFKA 660
Db 559 GDVDVNDWREHTKYNGYSLNHQVTHWFWKAVLMDSEKRIQLQVTCSTRVPMGFAE 618
QY 661 LOGSTGAAGPRFLT 675
Db 619 LYGSN---GPGSFTV 630

RESULT 9
T37900
probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T37900
R:Rieger, M.; McDougall, R.C.; Rajadream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z21752
A:Accession: T37900
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-671 <RIB>
A:CROSS-references: EMBL:AL117390; PIN:CA55856.1; GSPDB:GN00066; SPDB:SPAC1805.15C
A:Experimental source: strain 972h-; cosmid cl805
C:Genetics:
A:Gene: SPDB:SPAC1805.15C
A:Map position: 1
A:Insertions: 80/2; 105/1; 639/2
C:Superfamily: WW repeat homology
F:242-279/Domain: WW repeat homology <WWR>

Query Match 26.4%; Score 1026.5; DB 2; Length 671;
Best Local Similarity 32.9%; Pred. No. 1.3e-66;
Matches 248; Conservative 128; Mismatches 253; Indels 125; Gaps 18;

QY 6 KIRLTIVLCAKMLAKDFFRLDPPEAKIVVDGSGGQCHSDTIVKNTLDPKWCHDYLYVGKT 65
Db 7 EQLVTILHVEGLWKNGLRLSLKPYLLISVD-DDQFIKTNVAGSLRLSWGTQKLTVSPQ 65

QY 66 DSITISVWNKKHKKQAGLGCVRLLSNA---ISRLKDTGYQRLDCLKNPDDTDAVR 122
Db 66 STILLQLQFDEKQ-KNETSDGFGVLGAADVNSFLPNNPKDDYKTRITL---RSPSGSYR 120

QY 123 GQIVVSLQTRDRIGTGGSVDCRGLLENEGTVEDSGRPLSCFMEBPAP-----YTD 176
Db 121 GSWCLFKR-----SKFLPEELPADKSIQCTD 147

QY 177 STGAAG-----GCNCRFVRSPP---SQDORLQAOQLRNPVDRGSLQTPQNRPHGHS 225
Db 148 IIDDSAGGAWETRIDEFQHVYLYKSPQLSVISASISHEKLEN-----LTFK----- 192

QY 226 PELPEGYQRTTQGVYVFLHTGTGVTW-----HDPRIPLDNLNSVNCDEL 271
Db 193 -QLKEVFSG-----FLFNQSKSLKINLEYKVIKHLLEHYF-LALSVRQOQVAVEK 241

QY 272 GPLPGWEVRSTVSGRIYFVDHNNRTTQDTPRLHIMNHQCLKEPSQPLPLPSGSL 331
Db 242 GPLPAGWENRLESDYHVIYFVDSHTKTTWSDPRDNY-----ASDSVS 284

QY 332 DEELPAQRYERDLVQKLKVL--RHLSLQOQPAQHCRTEVSREEIFESYROIKMKRPD 389
```

```
Db 295 ENTDSIQINDEYQKIAMVDRPEMAVNDQA---LQIKVSRATTFEDAYDIISKLSVSD 341
Qy 390 LKKRLMVFRGEGDYGVAEWLYLICHEMLNPPYGLFOYSTDNIMQLNPDSINP 449
Db 342 MKKLLIRPNEDGLDYGVSSEFFYLISHAIFNPGYSLFYATDDNYGLQISPLSSVNP 401
Qy 450 DHLSYFHFVGRIMGLAVPHGHYINGGFTVPFYKQLGRPIQLSDLESVDPELHKSILWIL 509
Db 402 DFRSYFRFVGRVYGLAIYHRRYLDVQFLPFFYKRIQLKPLCLEDKVDYDEYVESLKWIK 461
Qy 510 ENDITPVLDDHTECVENHAFGRILQHELKPNGRNVPVTENKEKYYVRLVYNNRFRMGIEAQ 569
Db 462 NNDVDESCLNFSVENRFGESVTDLPNGRNIAVNNQNNYLLKALTEHKLVTSTBEQ 521
Qy 570 FLALQGFNELLIPQHLKPFQDKLELIGGLDKIDLNDWKSNTLKHCVADSNIVRWFV 629
Db 522 FNALAGGLNELIPDSVLQIFENELDTLNGKRDIDVDQDKRFTDYRSYTTDDIVIFW 581
Qy 630 QAVENFDEERRARLLQFVTSRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLPAKH 689
Db 582 EULSWSPBKAKLLQFATGSRLLPSLGFQWVGSD---GPKRFTIEKV-GHISQLPAKH 637
Qy 690 TCFNRIDIPPYESYKLYEKLKTAVEETCGFAVE 723
Db 638 TCFNRLDIPPYNSKKELEKLTIAICETAGFGE 671

RESULT 10
T49799
N:Alternate names: protein [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49799
R:Schulte, U.; Aign, V.; Hohenisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49799
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3839 <SCH>
A:Cross-references: EMBL:AL356834; GSPDB:GN00116; NCSP:B11B22.10
A:Experimental source: BAC clone B11B22; strain OR74A
C:Genetics:
A:Gene: NCSP:B11B22.10
A:Map position: 6
A:Intons: 16/3; 2607/1; 2623/1; 2658/1; 2987/2; 3204/3; 3694/1; 3809/1

Query Match 22.0%; Score 854.5; DB 2; Length 3839;
Best Local Similarity 39.4%; Pred. No. 6.6e-53;
Matches 178; Conservative 92; Mismatches 155; Indels 27; Gaps 8;

Qy 275 PRCWEVRSTVSGRIY-FVDHNRRTTQTPRLHHIMHQCQLKEPSQPLPLPSEGSLEDE 333
Db 3410 PP-----EDRIAGLFFTFTEBHRI-----LNELVHNPMLKMSGTFSLLVKNPKVLEFD 3458
Qy 334 ELPAQYERDLVQKLKVLREHLSLQOQAGHCRIEVSREEIFEESYRQIMKMRPKDLK-K 392
Db 3459 N-KRNVFNRSVHSKYQOTRHSPPPLQ-----LQVRREHVFHDSFRSLYYKAKDELKFG 3510
Qy 393 RLMVKFRGEGLDYGVAEWLYLICHEMLNPPYGLFOYSTDNIMQLNPDSINPDLH 452
Db 3511 KLNIRFGGEGVDAGGVTRFEWQVLSRQMPDNVYLFVFPVSSDRITTFHKNKLSPINDEHL 3570
Qy 453 SYFHFVGRIMGLAVPHGHYINGGFTVPFYKQLGRPIQLSDLESVDPELHKSILWILEND 512
Db 3571 PPFKFTGRIIGKALYGRLLCEYFSRAVYKRIIGKVPVKDMESDPDYKSLVWMLND 3630
Qy 513 ITPVLDDHTECVENHAFGRILQHELKPNGRNVPVTENKEKYYVRLVYNNRFRMGIEAQFLA 572
Db 3631 ITDIIETESVEDDFGEVKVWDLIENGRNIPVTENKREYVRLIVEHKLITSVKDQMK 3690
```

```
Qy 573 LQKGFNELIPQHLKPFQDKLELIGGLDKIDLNDWKSNTLKHCVADSNIVRWFV 632
Db 3691 FLTGFEHIIPEELIAIFNEQELLELSGLPDIDIDDKANTYHNSAGAPQIQWFRAV 3750
Qy 633 ETTDEERRARLLQFVTSRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLPAKHTCF 692
Db 3751 RSPDKKEELAKLQFVGTGSKVPLNGFKLEGMGVS---RFNIHRDYGSKORLPSSHTCF 3807
Qy 693 NRIDIPPYESYKLYEKLKTAVEETCGFA 721
Db 3808 NOLDPEYENYETLRSQLKAITAGSDYFGFA 3839

RESULT 11
T37964
probable ubiquitin ligase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T37964
R:Devlin, K.; Churcher, C.M.; Wood, V.; Bartelli, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21757
A:Accession: T37964
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-3227 <DEV>
A:Cross-references: EMBL:Z99531; PIDN:CAB16714.1; GSPDB:GN00066; SPDB:SPAC19D5.04
A:Experimental source: strain 972h-; cosmid C19D5
C:Genetics:
A:Gene: SPDB:SPAC19D5.04
A:Map position: 1

Query Match 21.8%; Score 848; DB 2; Length 3227;
Best Local Similarity 43.1%; Pred. No. 1.5e-52;
Matches 170; Conservative 79; Mismatches 125; Indels 20; Gaps 5;

Qy 344 LVQKLKVLK-----HLSLQOQAGHCRIEVSREEIFEESYRQIMKMRPKDLK 391
Db 2838 LVKNPKVLEFENKENYFNRLHHEAAKE-QYPLNITVREDHVFLOSYRALHFKDADEVK 2896
Qy 392 -KRLWVKFRGEGLDYGVAEWLYLICHEMLNPPYGLFOYSTDNIMQLNPDSINPD 450
Db 2897 FSKLNIHFRDEEGVDAGGVTRWLOVLARQVFPDYALFLPVTGDATTFHPNRDSSVNP 2956
Qy 451 HLSYFHFVGRIMGLAVPHGHYINGGFTVPFYKQLGRPIQLSDLESVDPELHKSILWILE 510
Db 2957 HLSFFKTRIIGKALYGRLLDCHFSRAVYKMLHRSVSVKDIESTLDPDYKSLVWMLN 3016
Qy 511 NDITPVLDDHTECVENHAFGRILQHELKPNGRNVPVTENKEKYYVRLVYNNRFRMGIEAQ 570
Db 3017 NDITDIITEFAVEKDVFGKTVVVDLIPNGRNIPVTSLNKQNVYNNRMDYIKLRESVKDQL 3076
Qy 571 LALQKGFNELIPQHLKPFQDKLELIGGLDKIDLNDWKSNTLKHCVADSNIVRWFV 630
Db 3077 KSLLDGFSDIIPSHLIQIFNEQELELLISGLPEIDIDDWKNTEYHGYNVSSPQVQFWFR 3136
Qy 631 AVETFDEERRARLLQFVTSRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLPAKHT 690
Db 3137 AVRSFDEERAKLQFATGTSKVPNGFKLEGMGSGF---QRFNIHKSYSGLNRLPQSH 3193
Qy 691 CFNRIDIPPYESYKLYEKLKTAVEETC---GPA 721
Db 3194 CFNQLDPEVDITYQLRSMULTAINESGEGFGFA 3227

RESULT 12
H96599
protein F14J16.10 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96599
R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Corn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
```

ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C:Authors: Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 Ker, M.; Wu, D.; Xu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A6141; MUID:21016719; PMID:11130712  
 A:Accession: H96599  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-4056 <SPO>  
 A:Cross-references: GB:A005173; NID:98778329; PIDN:AAF79338.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: F14J16.10  
 A:Map position: 1

Query Match 20.2%; Score 784; DB 2; Length 4056;  
 Best Local Similarity 39.4%; Pred. No. 9.8e-48;  
 Matches 172; Conservative 65; Mismatches 139; Indels 62; Gaps 7;  
 A:Map position: 1

QY 343 DLVQKLVRLHLSLQPPQ--AGHCRIEVSREEIFEESYQIMKMRPKDKKRLMKVFRG 400  
 DB 3624 DFNKKAYFRSRIHQDHISGLPRISVRRAYVLEDSYNQLRMRSPODLKGRNLVQFQG 3683  
 QY 401 BEGLDYGVARWMLYLLCHEMLNPPYGLFOYSTDNIMQLINPDSSINPDHLSYFHFVGR 460  
 DB 3684 BEGDAGLTREWYQLLSRVIFDKGALLFT-TVGNDAFQFNPNNSVYQTEHLSYFKFVGR 3742  
 QY 461 IMGLAVFHGHYINGGFTVPFYKQLLKPKIQLSDLESVDPELHKLWLTENDITPVLDT 520  
 DB 3743 MVAKALPDGQLLDVYFTRSPYKHLGVKVTYHDIEAVDPDYKXKLWLENDVSDILDT 3802  
 QY 521 F-----CQE-----HNAFORILQHLKPNGRN 542  
 DB 3803 FMSDADEKHILYKTEVRLMCFCLFFWCFIPKCHCIEIILISLMKKVTDYELKPGGRN 3862  
 QY 543 VPTVEENKKEYVLYVNWFRMRGEAQLALQKGFNELIPOHLKPPDKOKELIIGGLD 602  
 DB 3863 IRVTEETKHEYVDLVAGHILTNAEPQINAFLEGFNELIPRELVSINDELKELLSGLP 3922  
 QY 603 KID-----LNDWKNTRLKHCVADSNIVRWFQAVETFDERRARLLQFV 647  
 DB 3923 EIDCKLSDIQYLCAYPLDIDLKANTETSYTAGSPVIRWFWEVVKAFSKEDMARFLQFV 3982  
 QY 648 TGSTRVPLQGFALQGGTGAAGPRLFTHLIDANTDNLKPAHTCFNIDIPPYESYEKLY 707  
 DB 3983 TGTSKVPLEGFKALQ---GISGPORLQHKAYGAPERLPSAHTCFNQLDPEYQSKQEQ 4039  
 QY 708 EKLTAVEETC---GFA 721  
 DB 4040 ERLLLATHEASEGGFGFA 4056

RESULT 13  
 T01491  
 ubiquitin-protein ligase homolog F1707.15 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 04-Mar-2000  
 C:Accession: T01491  
 R:Vysotskaja, V.S.; Schwartz, J.R.; Toriumi, M.; Yu, G.; Kwan, A.; Oji, O.; Liu, S.; Li,  
 xtz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N.  
 submitted to the EMBL Data Library, June 1998  
 A:Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.  
 A:Reference number: Z14334  
 A:Accession: T01491  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1126 <YS>  
 A:Cross-references: EMBL:AC003671; NID:G2833627; PID:G3176690; GSPDB:GN00059; ATSP:F1707  
 A:Experimental source: cultivar Columbia  
 C:Genetics:

A:Gene: ATSP:F1707.15  
 A:Map position: 1  
 A:Introns: 118/3; 292/3; 373/3; 457/2; 826/3; 875/3; 902/3; 1057/3; 1093/2  
 C:Superfamily: ubiquitin-protein ligase homology  
 F:756-1120/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 20.1%; Score 780.5; DB 2; Length 1126;  
 Best Local Similarity 40.1%; Pred. No. 2.4e-48;  
 Matches 170; Conservative 67; Mismatches 138; Indels 49; Gaps 6;  
 A:Map position: 1

QY 343 DLVQKLVRLHLSLQPPQ--AGHCRIEVSREEIFEESYQIMKMRPKDKKRLMKVFRG 400  
 DB 707 DFNKKAYFRSRIHQDHISGLPRISVRRAYVLEDSYNQLRMRSPODLKGRNLVQFQG 766  
 QY 401 BEGLDYGVARWMLYLLCHEMLNPPYGLFOYSTDNIMQLINPDSSINPDHLSYFHFVGR 460  
 DB 767 BEGDAGLTREWYQLLSRVIFDKGALLFT-TVGNDAFQFNPNNSVYQTEHLSYFKFVGR 825  
 QY 461 IMGLAVFHGHYINGGFTVPFYKQLLKPKIQLSDLESVDPELHKLWLTENDITPVLDT 520  
 DB 826 MVAKALPDGQLLDVYFTRSPYKHLGVKVTYHDIEAVDPDYKXKLWLENDVSDILDT 885  
 QY 521 FCVEHNAFGRIL-----QHLEKPNGRNVPVTEENKKEYVLYVNWFRMRGEIAQFIAL 573  
 DB 886 FMSDADEKHILYKTEVTDYELKPGGRNIRVTEETKHEYVDLVADHILTSAIRPQINAF 945  
 QY 574 QKGFNELIPQHLKPPDKOKELIIGGLDKID----- 605  
 DB 946 BEGLNELIPRELVSINDELKELLSGLPEIDCKLAFYSFLHLFAYSFKVITILLSV 1005  
 QY 606 -----LNDWKNTRLKHCVADSNIVRWFQAVETFDERRARLLQFVGTSTRVPLQGFKA 660  
 DB 1006 PFCFLVDLLKANTETSYTAGSPVIRWFWEVVKAFSKEDMARFLQFVTGTSKVPLEGFK 1065  
 QY 661 LOGSTGAAGPRLFTHLIDANTDNLKPAHTCFNIDIPPYESYEKLYEKLTAVEET--- 717  
 DB 1066 LQ---GISGPORLQHKAYGAPERLPSAHTCFNQLDPEYQSKQVQERLLLAHEANEG 1122  
 QY 718 CGFA 721  
 DB 1123 FGFA 1126

RESULT 14  
 S69625  
 hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Mar-2001  
 C:Accession: S69625  
 R:Dietrich, F.S.  
 submitted to the EMBL Data Library, August 1995  
 A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.  
 A:Reference number: S69554  
 A:Accession: S69625  
 A:Molecule type: DNA  
 A:Residues: 1-3268 <DIE>  
 A:Cross-references: EMBL:U3050; NID:G927726; PIDN:AAB64910.1; PID:G927738; MIPS:YDR457w  
 C:Genetics:  
 A:Gene: SGD:TOM1  
 A:Cross-references: SGD:S0002865; MIPS:YDR457w  
 A:Map position: 4R

Query Match 19.8%; Score 769; DB 2; Length 3268;  
 Best Local Similarity 43.3%; Pred. No. 8.7e-47;  
 Matches 152; Conservative 70; Mismatches 125; Indels 4; Gaps 2;  
 A:Map position: 4R

QY 367 EVSREEIFEESYQIMKMRPKDLK-KELMVKFRGEGLDYGVARWMLYLLCHEMLNPPY 425  
 DB 2913 ITVREQVFLDSYRALFFKTNDKSKLEITFKGESGVADAGVYTRWYQVLSRQMFND 2972  
 QY 426 YGLFOYSTDNIMQLINPDSSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQL 485  
 DB 2973 YALFLVPSDKTTTFRNRTSGINPEHLSFFKFGIMIGKARDCQFLDCHESREVYNIL 3032

```
QY 486 GKPIQLSDLESVDPELHKLWILENDITPVLDHTFCVEHNAFGRILQHELKPNGRNVPV 545
Db 3033 GRPVSLEKDMESLDPYKSLWILENDITDIETTFVETDYDGHKVINIEGKDIIV 3092
QY 546 TEENKKEYVRLVYVNRFRGIEAQLALQKGFNELIPQHLKPPDQKELELIIGGLDKID 605
Db 3093 TEANKQDYVKKVEKYLQTSVKEQMDNPLVGFYALISKDLITIPDEQELELLISGLPDID 3152
QY 606 LNDWKSNTLKHCVKADSNIVRWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQGST 665
Db 3153 VDDWKNNTIYVNTATCKEVSFWAVRSFDEERAKLLQFVTGSKVPLNGFKELSGVN 3212
QY 666 GAAGRPLFTIHLIDANTONLPKXHCENRIDIPPESEKYLEKLLTAVEE 716
Db 3213 GVC---KFSIHRDFGSSERLPSHICFNQNLNPPYESIETLRGSLLLAINE 3260

RESULT 15
T20274
Hypothetical protein F45H7.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Aug-2000
C:Accession: T20274; T22257
R:Percy, C.
Submitted to the EMBL Data Library, November 1994
A:Reference number: Z19245
A:Accession: T20274
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-889 <W1>
A:Cross-references: EMBL:Z46793; PIDN:CAA86773.1; GSPDB:GNO0021; CESP:F45H7.6
A:Experimental source: clone C56G7
R:Percy, C.
Submitted to the EMBL Data Library, June 1994
A:Reference number: Z19538
A:Accession: T22257
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-889 <W12>
A:Cross-references: EMBL:Z34800; PIDN:CAA84325.1; GSPDB:GNO0021; CESP:F45H7.6
A:Experimental source: clone F45H7
C:Genetics:
A:Gene: CESP:F45H7.6
A:Map position: 3
A:Insertions: 30/3; 54/1; 118/1; 185/2; 213/1; 260/3; 288/2; 353/2; 454/1; 559/2; 642/2; 74
C:Superfamily: WW repeat homology
F:229-266/Domain: WW repeat homology <W1>
F:372-408/Domain: WW repeat homology <W2>

Query Match 17.2%; Score 669; DB 2; Length 889;
Best Local Similarity 28.7%; Pred. No. 2.3e-40;
Matches 168; Conservative 104; Mismatches 203; Indels 110; Gaps 15;

QY 232 YECRTTVQGVYFLHTQTGVSTWHDPRIPRDLN-----SWNCDELGLPLPP 276
Db 322 YENQDAMQ-----IYNESVVRHAIHQXLDLPPSKFENQPLFVFNLPADITQPLPS 376
QY 277 GWEVSTVSGRIYFVDHNNRTTQTPRLHHMHQCLKPSQPLPLPSEGS----- 329
Db 377 GWEC-ITMNRITVFLNHANKETSFYDPRIRF-----ETKTSRRGRSVPSRSSTAHKGI 430
QY 330 -----LEDELP--NQRVERDIV----- 345
Db 431 DHALISKCDLRKIAQDNFPQIAERISKMLLIERFGLAVASLANDLITIALSMLDSN 490
QY 346 -----QKLKVLRLHLSLQPPQAGHCRI--EVSREEIFESYQIMKMRPKDLKK-RL 394
Db 491 TEXLAGEGDNIMKFYEDMKKEKLGKPSRLCWKVSRRDLLDDAFRIILNVDPVLKKSRL 550
QY 395 MVKPRGEEGLDYGVARWELYLICHEMLNPYGLFQYSTNTIYMLQINPDSSINPDHLSY 454
Db 551 HIRFEGELALDYGGLGREFFILLRSLFHPKNGYFEY- GNDYHLQLRPRGCEETEKEKKW 609
```

```
QY 455 FHFVGRIMGLAVFHGHVINGGFTVPFYKQILGKPIQLSDLESVDPELHKLWILENDIT 514
Db 610 LILCGRVIALAVIHRCYIIDVFFINVIYKSLQKRPVTLMDPKESDAEFYKSMWLENDVV 669
QY 515 FVLDTFCVEHNAFGRI-----LQHELKPNGRNVPVTEENKKEYVRLVYN 559
Db 670 D-LEMSFVYSSVMNGKVSDDLTLSSIVSVALQAEQLVPGESQMVTEANKAEFIDLMCQ 728
QY 560 WRFRMGIEAQLALQKGFNELIPQHLKPPDQKELELIIGGLDKIDLNDWKSNTLKHCV 619
Db 729 KKAIRGVKEKPLEILLTSFNQILNDNLNLSLSSDLKRIILSGSLELDLNDWRTWTIYGGY 788
QY 620 ADSNI--VRWFQAVETFEERRARLLQFVTGSTRVPLQGFKALQGSTGAAGPLFTIHLI 678
Db 789 SDCHIVVWEVWEVETMTNQERFDLLLFVTGSSVVPFEGFSALRGNEETS---KFCIEKW 845
QY 679 DANTDNLPKAHTCFNRIIDIPPESEKYLEKLLTAVEETCGFAVE 723
Db 846 -GDATSPRAHTCFNRLQLPSYNTKQQLKSKLQQAIVNGMSYSIE 889

Search completed: September 21, 2004, 07:48:25
Job time : 35.9986 secs
```

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:29:24 : Search time 16.7111 Seconds  
(without alignment)  
2952.800 Million cell updates/sec

Title: US-10-009-945-2

Perfect score: 3894

Sequence: 1 GSSIKIRLTIVLCAKNAKK.....EKLYEKLITAVEETCGFAVE 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	3855	99.3	757	1	SUFL_HUMAN	O9hce7	homo sapien
2	3548	91.3	731	1	SUFL_YENLA	O9pun2	xenopus lae
3	3249.5	83.7	619	1	SUFL_MOUSE	O9cun6	mus musculus
4	3001	77.3	748	1	SUFL_HUMAN	O9haa4	homo sapien
5	1492.5	38.4	767	1	PUB1_SCHPO	O92462	schizosacch
6	1377	35.5	809	1	RSP5_YEAST	P39940	saccharomyc
7	1318	33.9	864	1	ITCH_MOUSE	O8c863	mus musculus
8	1288.5	33.2	903	1	ITCH_HUMAN	O95102	homo sapien
9	1259	32.4	1000	1	NER4_HUMAN	P45934	homo sapien
10	1259	31.9	922	1	WP21_HUMAN	O9homo	homo sapien
11	1234	31.8	870	1	WP21_MOUSE	O9dbh0	mus musculus
12	1232	31.7	918	1	WP21_MOUSE	O8bzz3	mus musculus
13	1227.5	31.6	887	1	NER4_RAT	Q62940	rattus norv
14	1225	31.5	870	1	WP21_HUMAN	O00308	homo sapien
15	1225	31.5	887	1	NER4_MOUSE	P46935	mus musculus
16	627.5	16.2	210	1	URB1_RAT	P51593	rattus norv
17	534	13.7	875	1	UE3A_HUMAN	Q05086	homo sapien
18	483	12.4	885	1	UE3A_MOUSE	O08759	mus musculus
19	481.5	12.4	892	1	HUL4_YEAST	P49985	saccharomyc
20	449.5	11.6	1050	1	HER3_HUMAN	Q15034	homo sapien
21	358	9.2	1992	1	TRB_HUMAN	Q14669	homo sapien
22	354	9.1	910	1	HUL5_YEAST	P53119	saccharomyc
23	278.5	7.2	2799	1	EDD_HUMAN	O95071	homo sapien
24	277	7.1	920	1	EDD_RAT	Q62671	rattus norv
25	262.5	6.8	1483	1	UF04_YEAST	P32202	saccharomyc
26	257	6.6	1647	1	YDE1_SCHPO	Q10435	schizosacch
27	238.5	6.1	472	1	YAP1_MOUSE	P46938	mus musculus
28	238.5	6.1	1620	1	HD1_HUMAN	O9ult8	homo sapien
29	222	5.7	2895	1	HYD_DROME	P51592	drosophila
30	173.5	4.5	1277	1	AIF1_RAT	O88382	rattus norv
31	172.5	4.4	1455	1	AIF1_HUMAN	O86u18	homo sapien
32	172	4.4	1275	1	AIF1_MOUSE	Q9wvq1	mus musculus
33	162	4.2	448	1	YAP1_CHICK	P46936	gallus gall

#### ALIGNMENTS

##### RESULT 1

```

SUFL_HUMAN          STANDARD;          PRT;   757 AA.
ID   SUFL_HUMAN          O9HCE7; O9UUT8;
AC   O9HCE7; O75853; O9UUT8;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
DE   protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)
DE   (hSMURF1).
GN   SMURF1 OR KIAA1625.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
[1]
SEQUENCE OF 10-731 FROM N.A. (ISOFORM SHORT).
RX   MEDLINE=99385348; PubMed=10458166;
PA   Zhu H., Kavaak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT   "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
RT   pattern formation.";
RL   Nature 400:687-693(1999).
[2]
SEQUENCE OF 20-731 FROM N.A. (ISOFORM SHORT).
RA   Stoneking T., Bauer C., O'Neal D.;
RN   Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A. (ISOFORM LONG).
RC   TISSUE=Brain;
RX   MEDLINE=20450683; PubMed=10997877;
RA   Nagase T., Kikuno R., Nakayama M., Hirosewa M., Ohara O.;
RT   "Prediction of the coding sequences of unidentified human genes.
RT   XVIII. The complete sequences of 100 new cDNA clones from brain which
RT   code for large proteins in vitro.";
RL   DNA Res. 7:273-281(2000).
CC   -!- FUNCTION: Interacts with receptor-regulated SMADs specific for the
CC   bmp pathway (SMAD1 AND SMAD5) in order to trigger their
CC   ubiquitination and degradation and hence their inactivation.
CC   -!- ALTERNATIVE PRODUCTS:
CC   Event-Alternative splicing; Named isoforms=2;
CC   Name=Long;
CC   IsoId=O9HCE7-1; Sequence=Displayed;
CC   Name=Short;
CC   IsoId=O9HCE7-2; Sequence=VSP_006812;
CC   -!- SIMILARITY: Contains 1 C2 domain;
CC   -!- SIMILARITY: Contains 2 WW domains.
CC   -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC   domain.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

34 161.5 4.2 454 1 YAP1_HUMAN          P46937 homo sapien
35 143.5 3.7 383 1 SAVI_MOUSE          Q9h4b6 homo sapien
36 143.5 3.7 386 1 SAVI_MOUSE          Q8vdb2 mus musculus
37 128.5 3.3 1555 1 RIM2_RAT           Q9jls1 rattus norv
38 128 3.3 1813 1 UN13_CABEL        P27715 caenorhabdi
39 119.5 3.1 1530 1 RIM2_MOUSE        Q9egz7 mus musculus
40 119 3.1 575 1 BAG3_HUMAN        Q95817 homo sapien
41 118 3.0 1188 1 RIM2_HUMAN        Q9uc26 homo sapien
42 116.5 3.0 577 1 BAG3_MOUSE        Q9jlv1 mus musculus
43 115.5 3.0 1692 1 RIM1_HUMAN        Q86ur5 homo sapien
44 114.5 2.9 803 1 RSG5_HUMAN        C43374 homo sapien
45 114 2.9 593 1 CNE5_HUMAN        Q9hch3 homo sapien

```

```

CC -----
DR EMBL; AF199364; AAF08298.2; -
DR EMBL; AC004893; AAC62434.1; -
DR EMBL; AB046845; BAB13451.1; ALT_INIT.
DR HSP; Q13526; 1FN.
DR MIM; 605568; -
DR GO; GO:0005622; C:intracellular; TAS.
DR GO; GO:0000211; F:protein degradation tagging activity; IDA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
DR GO; GO:0030154; P:cell differentiation; IDA.
DR GO; GO:0007398; P:ectoderm development; TAS.
DR GO; GO:0030314; P:negative regulation of BMP signaling pathway; TAS.
DR GO; GO:0006512; P:ubiquitin cycle; IDA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2 CalB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2_1.
DR Pfam; PF00632; HECT_1.
DR Pfam; PF00397; WW_2.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW_2.
DR PROSITE; PS00499; C2 DOMAIN 1; 1.
DR PROSITE; PS00004; C2 DOMAIN_2; 1.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW DOMAIN 1; 1.
DR PROSITE; PS00020; WW DOMAIN_2; 2.
DR Ubl conjugation pathway; Ligase; Repeat; Alternative splicing.
FT DOMAIN 1 99 C2 DOMAIN.
FT DOMAIN 234 267 WW 1.
FT DOMAIN 306 339 WW 2.
FT BINDING 420 757 HECT.
FT BINDING 725 725 UBIQUITIN.
FT VARSPIC 269 294 Missing (in isoform short).
FT MUTAGEN 725 725 /FTID=VSP_006812.
FT SEQUENCE 757 AA; 86113 MW; 89A171CFC47B40E9 CRC64;
Query Match 99.3%; Score 3855; DB 1; Length 757;
Best Local Similarity 96.5%; Pred. No. 3.2e-274;
Matches 722; Conservative 0; Mismatches 0; Indels 26; Gaps 1;
QY 2 GSSIKIRLTVCANLAKKQFFRLPDPFAKIVDVGSGQCHSTDTVKNTLDPKKNQHYDLY 61
DB 10 GSSIKIRLTVCANLAKKQFFRLPDPFAKIVDVGSGQCHSTDTVKNTLDPKKNQHYDLY 69
QY 62 VGKTDSTISVNNHKKIKHGAGFLGCVRLLSNAISRLKDTGYQRLDLCKLNPSTDAV 121
DB 70 VGKTDSTISVNNHKKIKHGAGFLGCVRLLSNAISRLKDTGYQRLDLCKLNPSTDAV 129
QY 122 RGQIVSLQTRDRTGTGSGVDCRGLLENGTVYEDSGGRPLSCFEEBPAPYTDSTGAA 181
DB 130 RGQIVSLQTRDRTGTGSGVDCRGLLENGTVYEDSGGRPLSCFEEBPAPYTDSTGAA 189
QY 182 AGGNCNRFVSPSQDRLQRLNPVDRGSLQTPQNRPHGHSPELPEGYEQRITTVQGG 241
DB 190 AGGNCNRFVSPSQDRLQRLNPVDRGSLQTPQNRPHGHSPELPEGYEQRITTVQGG 249
QY 242 VYFLHTGTGVTWHDRI - - - - - PRLNSVNCDELGLP 275
DB 250 VYFLHTGTGVTWHDRI - - - - - PRLNSVNCDELGLP 309
QY 276 PGWEVRSTVSGRIYFVDHNRRTQFTDPRLHNMHQCQLKPSQPLPSEGSLEDEEL 335
DB 310 PGWEVRSTVSGRIYFVDHNRRTQFTDPRLHNMHQCQLKPSQPLPSEGSLEDEEL 369
QY 336 PAQRYEDLVQKLVLRHLSLQOPQACHRIEVSREIEFESYRQIMKRPDLKKRLM 395
DB 370 PAQRYEDLVQKLVLRHLSLQOPQACHRIEVSREIEFESYRQIMKRPDLKKRLM 429
QY 396 VKFRGEGLDYGAVREWLKYLCHMLNPPYGLFOYSTDNIMYQLINPDSSINPDHLSYF 455

```

```

DB 430 VKFRGEGLDYGAVREWLKYLCHMLNPPYGLFOYSTDNIMYQLINPDSSINPDHLSYF 489
QY 456 HFVGRIMGLAVFHGHYINGGFTVPFYPKOLLGKPIQLSDLESVDPLHKLKSLVILENDITP 515
DB 490 HFVGRIMGLAVFHGHYINGGFTVPFYPKOLLGKPIQLSDLESVDPLHKLKSLVILENDITP 549
QY 516 VLDHTFCVHNFAFRILQHELKPNRNPVTEENKKEYVRLYVNWRFMRGTEAFLALQK 575
DB 550 VLDHTFCVHNFAFRILQHELKPNRNPVTEENKKEYVRLYVNWRFMRGTEAFLALQK 609
QY 576 GSENELIPHLLKPPDKQKLELIIGLDKIDLDNKSNTLKHCVADSNIVFWFQAVET 635
DB 610 GSENELIPHLLKPPDKQKLELIIGLDKIDLDNKSNTLKHCVADSNIVFWFQAVET 669
QY 636 DEERARLLQFVTGSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTDNLKHAHTCFNRI 695
DB 670 DEERARLLQFVTGSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTDNLKHAHTCFNRI 729
QY 696 DIPPYESYEKLYEKLITAVEETCGFAVE 723
DB 730 DIPPYESYEKLYEKLITAVEETCGFAVE 757

RESULT 2
ID SUPL_XENLA STANDARD; PRT; 731 AA.
AC Q9PUN2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--
GN SMURF1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Blascula;
RX MEDLINE=99385348; PubMed=10458166;
RT Zhu H., Kaysak P., Abdollah S., Wrana J.L., Thomsen G.H.;
RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic
RT pattern formation.";
RL Nature 400:687-693(1999).
CC -!- FUNCTION: INTERACTS WITH RECEPTOR-REGULATED SMADS SPECIFIC FOR THE
CC BMP PATHWAY (SMAD1 AND SMAD5) IN ORDER TO TRIGGER THEIR
CC UBIQUITINATION AND DEGRADATION AND HENCE THEIR INACTIVATION. MAY
CC REGULATE ECTODERMAL DIFFERENTIATION AND PATTERN BY MODULATING BMP
CC SIGNALING AND MAY ENHANCE CELLULAR RESPONSIVENESS TO THE SMAD2
CC (ACTIVIN/TGF-BETA) PATHWAY.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM THE EGG STAGE TO THE SWIMMING
CC TADPOLE, WITH MAXIMUM LEVELS OBSERVED IN THE STAGES FROM EGG TO
CC GASTRULA. AT GASTRULATION DISTRIBUTED UNIFORMLY IN EMBRYONIC
CC ECTODERM AND INVOLUTING MESODERM, AND EXPRESSION GRADUALLY
CC LOCALIZES TO THE NERVOUS SYSTEM. AT EARLY TADPOLE STAGES EXPRESSED
CC IN THE CNS, EYE, BRANCHIAL ARCHES, KIDNEY AND SOMITES.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 2 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF169310; AAD52564.1; -
CC GO; GO:0005622; C:intracellular; TAS.

```



modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AK015264; BAB29770.2; ALT\_INIT.  
 HSP; Q13526; 1PIN.  
 MGD; MGI1923038; 4930431E10Rik.  
 GO; GO:0005622; C:intracellular; ISS.  
 GO; GO:000211; F:protein degradation tagging activity; ISS.  
 GO; GO:0004842; F:ubiquitin-protein ligase activity; ISS.  
 GO; GO:0030154; P:cell differentiation; ISS.  
 GO; GO:0007398; P:ectoderm development; ISS.  
 GO; GO:0030514; P:negative regulation of BMP signaling pathway; ISS.  
 GO; GO:0006512; P:ubiquitin cycle; ISS.  
 InterPro; IPR000008; C2.  
 InterPro; IPR000569; HECT domain.  
 InterPro; IPR001202; WW\_RspS\_WMP.  
 Pfam; PF00632; HECT; 1.  
 Pfam; PF00397; WW; 2.  
 SMART; SM00119; HECTC; 1.  
 SMART; SM00456; WW; 2.  
 PROSITE; PS00499; C2\_DOMAIN\_1; PARTIAL.  
 PROSITE; PS00004; C2\_DOMAIN\_2; PARTIAL.  
 PROSITE; PS0237; HECT; 1.  
 PROSITE; PS01159; WW\_DOMAIN\_1; 1.  
 PROSITE; PS0020; WW\_DOMAIN\_2; 2.  
 Ubl conjugation pathway; Ligase; Repeat.  
 NON\_TER 1  
 DOMAIN 125 158 WW 1.  
 DOMAIN 171 204 WW 2.  
 DOMAIN 285 619 HECT.  
 DOMAIN 35 38 POLY-GLY.  
 SEQUENCE 619 AA; 70902 MW; 4CBZF8624A7B525 CRC64;

Query Match 83.7%; Score 3249.5; DB 1; Length 619;  
 Best Local Similarity 97.3%; Pred. No. 4.9e-230;  
 Matches 605; Conservative 6; Mismatches 8; Indels 3; Gaps 1;

102 DTGQRLDLCKLNPSDSDAVERGQVWVSLQTRDRIGTGGVSDRCGLLENBGTVYDSGPG 161  
 1 DTGQRLDLCKLNPSDSDAVERGQVWVSLQTRDRIGGGSVSDRCGLLENBGTVYDSGPG 60

162 RPLSCFEEAPYDSTDAAGAGGNCRFVSPDQRLQORLANPVRGSLTPQNRPH 221  
 61 RPLSCLMEAPYDGTGAAGGNCRFVSPQNRLLVQRLQNPVEVQGLTPQNRPH 120

222 GHSQPELPEGEQRTTVGGQVYFHTGTGVTWHDPRIPDLNSVNCDELGLPDPGWEVR 281  
 121 GHSQPELPEGEQRTTVGGQVYFHTGTGVTWHDPRIPDLNSVNCDELGLPDPGWEVR 180

282 STVSGRIYFVDHNRRTQFTDPRLLHMHQCLKEPSQPLPSPSGSDEDELPQRYE 341  
 181 STVSGRIYFVDHNRRTQFTDPRLLHMHQCLKEPSQPLPSPSGSDEDELPQRYE 240

342 RDLVQKLVLBEHLSLQOPQAGHCRIVSEEEFEESVYRIMKWRKDLAKRLMKVFRGE 401  
 241 RDLVQKLVLBEHLSLQOPQAGHCRIVSEEEFEESVYRIMKWRKDLAKRLMKVFRGE 300

402 EGLDYGGVAREWLYLICHEMLNPPYGLFQYSTDNIMYLQINPDSSINPDHLSYFHVGR 461  
 301 EGLDYGGVAREWLYLICHEMLNPPYGLFQYSTDNIMYLQINPDSSINPDHLSYFHVGR 360

462 MGLAVFHGHYNGGFTVPFYKQLIGKPIQLSDLESVDPELHKSILVILENDITPVLDTHT 521  
 361 MGLAVFHGHYNGGFTVPFYKQLIGKPIQLSDLESVDPELHKSILVILENDITPVLDTHT 420

522 CVEHNAFRIILQHELKPNRNPVTEENKKEYVRLVYNWRFMRGIEAQFALQKGFNEL 581  
 421 CVEHNAFRIILQHELKPNRNPVTEENKKEYVRLVYNWRFMRGIEAQFALQKGFNEL 480

582 PQHLKPPDQKLELITGGLDKIDLNDKSNTRLKHCVADSNIVRWQAVETFDERRA 641  
 481 PQHLKPPDQKLELITGGLDKIDLNDKSNTRLKHCVADSNIVRWQAVETFDERRA 540

QY 642 RLLOFVTGSTRVPLQKALOGSTGAAGPRLFTIHLIDANTDNLKPAHTCFNRIDIPPE 701  
 DB 541 RLLOFVTGSTRVPLQKALQ---GAAGPRLFTIHLIDANTDNLKPAHTCFNRIDIPPE 597

QY 702 SYEKLYEKLTAVEETCGFAVE 723  
 DB 598 SYEKLYEKLTAVEETCGFAVE 619

RESULT 4  
 SUP2 HUMAN  
 ID SUP2 HUMAN STANDARD; PRT; 748 AA.  
 AC Q9HAU4; Q9H260;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Smad ubiquitination regulatory factor 2 (EC 6.3.2.-) (Ubiquitin--  
 protein ligase SMURF2) (Smad-specific E3 ubiquitin ligase 2)  
 DE (hSMURF2).  
 GN SMURF2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF 251-PRO--VAL-284 AND  
 RP 297-GLY--LEU-330.  
 RX PubMed=1163210;  
 RA Kavsak P., Rasmussen R.K., Causing C.G., Bonni S., Zhu H.,  
 RA Thomsen G.H., Wrana J.L.;  
 RT "Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the  
 RT TGF-beta receptor for degradation.";  
 RN Mol. Cell 6:1365-1375(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF 251-PRO--VAL-284 AND CYS-716.  
 RP MEDLINE=20538422; PubMed=11015919;  
 RA Lin X., Liang M., Feng X.-H.;  
 RT "Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent  
 RT degradation of Smad2 in transforming growth factor-beta signaling.";  
 RN J. Biol. Chem. 275:36818-36822(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-716.  
 RP MEDLINE=21107656; PubMed=1158580;  
 RA Zhang Y., Chang C., Genling D.J., Hemmati-Briylanlou A., Derynck R.;  
 RT "Regulation of Smad degradation and activity by Smurf2, an E3  
 RT ubiquitin ligase.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 98:974-979(2001).  
 CC -!- FUNCTION: Interacts with SMAD1, SMAD2 and SMAD7 in order to  
 CC trigger their ubiquitination and proteasome-dependent degradation.  
 CC Enhances the inhibitory activity of SMAD7 and reduces the  
 CC transcriptional activity of SMAD2. Coexpression of SMURF2 with  
 CC SMAD1 results in considerable decrease in steady-state level of  
 CC SMAD1 protein and a smaller decrease of SMAD2 level.  
 CC -!- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, SMAD6 and SMAD7 but  
 CC not SMAD4.  
 CC -!- SUBCELLULAR LOCATION: Nuclear. Cytoplasmic in the presence of  
 CC SMAD7.  
 CC -!- TISSUE SPECIFICITY: Widely expressed.  
 CC -!- DOMAIN: The second and third WW domains are responsible for  
 CC interaction with R-SMAD (SMAD1, SMAD2 and SMAD3).  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 3 WW domains.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC      EMBL; AF310676; AAG45422.1; --
DR      EMBL; AF301463; AAG25641.1; --
DR      EMBL; AY014180; AAG50421.1; --
DR      HSSP; Q13526; 1PIN.
DR      MIM; 605532; -.
DR      GO; GO:0004842; F-ubiquitin-protein ligase activity; NAS.
DR      GO; GO:0016481; P-negative regulation of transcription; NAS.
DR      GO; GO:0017015; P-regulation of TGF-beta receptor signaling pa...; NAS.
DR      GO; GO:0006511; P-ubiquitin-dependent protein catabolism; NAS.
DR      InterPro; IPR000008; C2.
DR      InterPro; IPR008973; C2_CalB.
DR      InterPro; IPR000569; HECT domain.
DR      InterPro; IPR001202; WW_Rsp_WWP.
DR      Pfam; PF00168; C2; 1.
DR      Pfam; PF00632; HECT; 1.
DR      Pfam; PF00397; WW; 3.
DR      SMART; SM00239; C2; 1.
DR      SMART; SM00119; HECTC; 1.
DR      SMART; SM00456; WW; 3.
DR      PROSITE; PS00499; C2_DOMAIN_1; 1.
DR      PROSITE; PS00004; C2_DOMAIN_2; 1.
DR      PROSITE; PS02337; HECT; 1.
DR      PROSITE; PS01159; WW_DOMAIN_1; 1.
DR      PROSITE; PS00020; WW_DOMAIN_2; 3.
KW      Ub1 conjugation pathway; Ligase; Repeat; Nuclear protein.
FT      DOMAIN 1 98
FT      DOMAIN 157 190
FT      DOMAIN 251 284
FT      DOMAIN 297 330
FT      DOMAIN 414 748
FT      MUTAGEN 251 284
FT      MUTAGEN 297 330
FT      MUTAGEN 716 716
FT      MUTAGEN 716 716
FT      MUTAGEN 716 716
FT      MUTAGEN 716 716
FT      CONFLICT 6 6
FT      CONFLICT 6 6
SQ      SEQUENCE 748 AA; 86195 MW; 3042B443A3755762 CRC64;

Query Match
Best Local Similarity 77.3%; Score 3001; DB 1; Length 748;
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

QY      5 IKRLTVLCANLAKDFFRLPDPFAKIVVDGSGQCHSTDTYKNTLDPKWNQHYDLYVGK 64
DB      12 VKRLTVLCANLAKDFFRLPDPFAKIVVDGSGQCHSTDTYKNTLDPKWNQHYDLYGK 71

QY      65 TDSITISVNNHKKIKQAGFLGCVRLLSNAISLKDGTQRLDCKLNPSDQDAVRGQ 124
DB      72 SDSVITISVNNHKKIKQAGFLGCVRLLSNAISLKDGTQRLDCKLGNPDNDTVRGQ 131

QY      125 IVVSLQTRDRIQTGGVSDCRGLLENE-----GTVY----- 155
DB      132 IVVSLQSRDRIQTGGVSDCRGLFNDLDPDQWEERTASGLIYLNHITRITQWERTP 191

QY      156 --EDSGPRPLSCFWEERAPYDSTDGAAGGNCRFVESPQDRLQRLNPDVRGSL 213
DB      192 ASEYSGRPLSCFVDENTPISTINGATCG-----QSSDPLAERRVRSORHNYM 242

QY      214 QTPQNRPHGQSPQLPEGVQRSTTVCGQVYFLHTGTGVSTWHDDRIPELNSVNCDELGP 273
DB      243 ----SRTHLHTPPDLPPEGVQRSTTVCGQVYFLHTGTGVSTWHDDRPVPLDNLNCEELGP 298

QY      274 LPPGWEVSTVSGRIYFVDHNNRTTQFTDPR-----LHHMNHQCOLKEPSPQLPLPSEGS 329
DB      299 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQLKQKQDQVQV-----S 354

QY      330 L---EDELPAQRYERDLVQKLKVLRLHLSLQDQPAQAGHCRIEVSEEFEEFESYQIMKMR 386

```

```

DB      355 LCPDDTECLATVPYKRDVQKLKILRQELSQQQPQAGHCRIEVSEEFEEFESYQIMKMR 414
QY      387 PKDLKRLMKVPRGEBGLDYGVARWLYLLCHEMLNPPYGLFOYSTDNTYMLQINPDSS 446
DB      415 PKDLKRLMKVPRGEBGLDYGVARWLYLLCHEMLNPPYGLFOYSRDDIYTLQINPDSA 474
QY      447 INPDHLSYFHFVGRIMGLAVFHGHYNGFTVPFYKQLLKPTQLSDLESVDDELHKSLV 506
DB      475 VNPBLSYFHFVGRIMGLAVFHGHYDGGFTLFFYKQLLGSITLDDMELVDPLHNSLV 534
QY      507 WILENDITPVLDTFCVEHNAFGRILQHELUKPNRNVPTTEENKKEYVRLYVNVNRMERI 566
DB      535 WILENDITGLVDHTFCVEHNAFGEIILQHELUKPNKGIPTVNEENKKEYVRLYVNVNRMERI 594
QY      567 EAQFLALQKGFNELIPQHLKLPDQKELELIIGSLDKLDLNDKSNTRLKHCVCVADSNIVR 626
DB      595 EAQFLALQKGFNEVIPHOLLKTFDEKELELIIGLGDIDVNDKVNTRLKHCCTPDNSIVK 654
QY      627 WFWQAVETDEERRARLLQFVTGSTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNL 686
DB      655 WFWKAVEFFDEERRARLLQFVTGSSRVPLQGFKALQ---GAAGPRLFTIHOIDACTNNLP 711
QY      687 KAHTCFNRIDIPPEYSEKLYEKLTAVERTCGFAYE 723
DB      712 KAHTCFNRIDIPPEYSEKLYEKLTAIBETCGFAVE 748

RESULT 5
PUB1_SCHPO STANDARD; PRT; 767 AA.
AC Q92462; O14454;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase pub1 (EC 6.3.2.-).
GN PUB1 OR SPAC1197.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
CX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=96205868; PubMed=8635463;
RA Nefsky B., Beach D.;
RT "Pub1 acts as an E3-AP-like protein ubiquitin ligase in the
RT degradation of cdc25."
RL EMBO J. 15:1301-1312(1996).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=J227;
RC MEDLINE=97340937; PubMed=9197411;
RA Saleki R., Jia Z., Karagiannis J., Young P.G.;
RT "Tolerance of low pH in Schizosaccharomycetes pombe requires a
RT functioning pub1 ubiquitin ligase."
RL Mol. Gen. Genet. 254:520-528(1997).
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

```





358 QOPQAGHRIEVESEBIFESYRQIMKMRPKDKKGLMVKFRGEBGLDYGGVAREWLYLL 417  
 449 -RILPGQCHIKVRKNIFEDAYQEIINROTPEDKKGLMIKFDGEEGLDYGGVAREWLYLL 507  
 418 CHEMLNPNYGLFOYSTNIMVLOINPDSSINPHLSYFHFVGRIMGLAVFHGHIYNGGRT 477  
 508 SHENFNPFCYFISAYDNTIQINNSGINPELNYFXFIVGVGLGVFHRFLDAFV 567  
 478 VPYKQLLGPQLSLESVDPELHSLVWILENDITPVLDHTFCVHEHNAFRILQHEUK 537  
 568 GALYKMLRKVVLQDMGVDAEVNSLWMLNSIDGVLDLTFFSADDERFGEVVTVDLK 627  
 538 PNGBNVPTENKEKYVRLVYVWFRGTEAOFALQKGFNELIPQHLKPPDOKELELI 597  
 628 PDGRNIEVTGNKKEYVELYQWRIVDVQEQKAFMDGFNEIIPDLVTVDERELELL 587  
 598 IGGDKIDLDLWKSNTLRKHCVADSNIVRFWQAVETFDERRARLLQFVTGSTRVPLQG 657  
 688 IGGIAEDIEDKXHTDYRGYQSDSEVIOFWKCVSEWNEQARLLQFTTGTSTRIPVNG 747  
 658 FKALQSTGAGPRLFTIHLIDANTDNLPAHCTCFNRIDIPPYSEYKLYEKLITAVEET 717  
 748 FKDLQSGD--GPRRFTIERA-GEVQQLPKSHTCFNRVLDLPQYVDYDSMKQKLTAVEET 803  
 718 CGRAVE 723  
 804 IGFQGE 809

RESULT 7  
 ID ITCH MOUSE STANDARD; PRT; 864 AA.  
 AC Q8C83; 054971; PRT; 864 AA.  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Itchy E3 ubiquitin protein ligase (EC 6.3.2.-).  
 GN ITCH  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND DISEASE.  
 RC STRAIN=C3H/HeJ; TISSUE=Kidney;  
 RX MEDLINE=98122574; PubMed=9462742;  
 RA Perry W.L., Hustad C.M., Swing D.A., O'Sullivan T.N., Jenkins N.A.,  
 RA Copeland N.G.;  
 RA "The itchy locus encodes a novel ubiquitin protein ligase that is  
 RT disrupted in a18H mice."  
 RL Nat. Genet. 18:143-146(1998).  
 [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nakaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Gramond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., McKinnon L., McKenzie L., Miki H.,  
 RA Nagashima T., Nomata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,

Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 [3]  
 RN FUNCTION, AND INTERACTION WITH JUN AND JUNE.  
 RP MEDLINE=21864584; PubMed=11828324;  
 RX Fang D., Elly C., Gao B., Fang N., Altman Y., Joazeiro C., Hunter T.,  
 RA Copeland N., Jenkins N., Liu Y.C.;  
 RA "Dysregulation of T lymphocyte function in itchy mice: a role for  
 RT Itch in Th2 differentiation."  
 RL Nat. Immunol. 3:281-287(2002).  
 [4]  
 RN INTERACTION WITH NOTCH1, AND MUTAGENESIS OF CYS-832.  
 RP MEDLINE=20549573; PubMed=10940313;  
 RX Qiu L., Joazeiro C., Fang N., Wang H.-Y., Elly C., Altman Y., Fang D.,  
 RA Hunter T., Liu Y.-C.;  
 RA "Recognition and ubiquitination of Notch by Itch, a hec-type E3  
 RT ubiquitin ligase."  
 RL J. Biol. Chem. 275:35734-35737(2000).  
 [5]  
 RN INTERACTION WITH OCNL.  
 RP MEDLINE=21895815; PubMed=11782481;  
 RX Traweger A., Fang D., Liu Y.-C., Stelzhammer W., Krizbai I.A.,  
 RA Fresser F., Bauer H.-C., Bauer H.;  
 RA "The tight junction-specific protein occludin is a functional target  
 RT of the E3 ubiquitin-protein ligase Itch."  
 RL J. Biol. Chem. 277:10201-10208(2002).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
 CC then directly transfers the ubiquitin to targeted substrates.  
 CC Regulates the transcriptional activity of several transcription of  
 CC factors, and probably plays an important role in the regulation of  
 CC immune response.  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Interacts via its WW domains with OCLN, NOTCH1, JUN and  
 CC JUNE. Interacts with DRPLA and NFE2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q8C863-1; Sequence=Displayed;  
 CC Note=Major form;  
 CC Name=2;  
 CC IsoId=Q8C863-2; Sequence=VSP\_008452, VSP\_008453;  
 CC -!- TISSUE SPECIFICITY: Widely expressed.  
 CC -!- PTM: Phosphorylated on tyrosine residues (By similarity).  
 CC -!- DISEASE: Defects in Itch are the cause of the itchy phenotype  
 CC which is an inflammatory and immunological condition characterized  
 CC by inflammation in the lung and stomach, hyperplasia in lymphoid  
 CC and hematopoietic cells and constant itching in the skin.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -!- SIMILARITY: Contains 4 WW domains.  
 CC -!- CAUTION: It is uncertain whether Met-1 or Met-11 is the initiator.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AF037454; AB99764.1; ALT\_INIT.

[illegible]

Qy	325	-----PSGSLDEDELPAQRVEDLVOKLKVLR--HLSLQQPQAGHCRIEVSRE	372
Dd	461	NRRATTYIDERTGKSALDNGFQAYVRDFRAKQVFRWCQLAMPO----HIKITVTRK	516
Qy	373	EIPESYRQIMKWMPDKLKRLMKVPRGEGLDYGGVAREWLYLLCHEMLNPPYGLFOYS	432
Dd	517	TLFDSFQQIIMSFPQDLRRRLWIIFPGEBGLDYGGVAREWFELSHVEVLNPWCUFEYA	576
Qy	433	TDNIYMLQINPDSSINDPHUSYFFHFVGRINGLAVFHGHYINGFTVPFYKQLLGKPTQLS	492
Dd	577	GKNYCILQINPASYPNDPHLKYPFRFGRTFAMALFKGFTDTGFSLPFYKRILNKPVGLK	636
Qy	493	DLESVDPELHKSLVWILENDITPV-LDHTPCFVEHNAFGRILQHLELKPENGNVPVTSENKK	551
Dd	637	DLESIDEPNSLIWYKNNIESCGLEMVPSVDKEILGKSHDLKPNNGNILUTEENKE	696
Qy	552	EYRVLYNVFMFGIRBAQFLALQGFNELIPQHLKPFDPQKELELIGGLDKIDLNDWKS	611
Dd	697	EYIRMWAEWLRSRGVBEQTQAFFEGENEILPQQYLQVFDAKELEVLLCGMQEIDLNDWR	756
Qy	612	NTRLKHCVADSNIVRMFWQAVETFDERRARILLQVGTSTRPVLOGFKALQGSGTGAAGR	671
Dd	757	HAIYRHVTRTSKOIMFWQFVKIDEKNRWLLQFVTGCRPLVGGVGFADLMGSN---GPQ	813
Qy	672	LFTIHLDANTDNLPKAHTCNRIDIPPYESYEKLKYELLTAVEETCGFAVE	723
Dd	814	KFCIEKV-GKENLWPRSHTCFNRDLDPFKYSYEQLKKEKLLFAETEGETGQE	864

RESULT 8

ID	ITCH_HUMAN	STANDARD;	PRT; 903 AA.
AC	Q96J02; Q43584; Q96F66; Q9BY75; Q9H451; Q9H4U5;		
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Itchy homolog E3 ubiquitin protein ligase (EC 6.3.2.-) (Itch)		
DE	(Atrophin-1-interacting protein 4) (AIP4) (NFE2-associated polypeptide 1) (NAPPI).		
GN	ITCH.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RP	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH NFE2.		
RC	TISSUE=Leukemia;		
RX	MEDLINE=21218930; PubMed=11318614;		
RA	Chen X., Wen S.-C., Fukuda M.N., Gavva N.R., Hsu D.-W., Akama T.O.,		
RA	Yang-Peng T.L., Shen C.K.J.;		
RT	"Human ITCH is a co-regulator of the hematopoietic transcription		
RL	factor NF-E2";		
RL	Genomics 73:238-241(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RA	Miyazaki K., Okamoto Y., Sakamoto M., Nakagawara A.;		
RT	"Homo sapiens mRNA for ubiquitin protein ligase Itch, complete cds.";		
RL	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.		
RP	[3]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21638749; PubMed=11780052;		
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,		
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,		
RA	Bailley J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,		
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,		
RA	Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Clee C.M.,		
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,		
RA	Clegg S., Cobley V.E., Collier R.E., Connor R., Corby N.R.,		
RA	Coulson A., Coville G.J., Deadman R., Dhani P., Dunn M.,		
RA	Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,		
RA	Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,		
RA	Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,		

RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knighes A., Laird G.K., Lawlor S.,  
 RA Lehtvala M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,  
 RA Milne S., Mistry D., Moore M.J.P., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A.C., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.,  
 RT "The DNA sequence and comparative analysis of human chromosome 20."  
 RL Nature 414:865-871 (2001).  
 RN [4]  
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RP TISSUE=Kidney, and placenta;  
 RC MEDLINE=22398257; PubMed=12477932;  
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [5]  
 RN SEQUENCE OF 83-303 FROM N.A. (ISOFORM 2), AND INTERACTION WITH DRPLA.  
 RP TISSUE=fetal brain;  
 RC MEDLINE=98313405; PubMed=9647693;  
 RX Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,  
 RA Kaminsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.;  
 RT "Atrophin-1, the DRPLA gene product, interacts with two families of WW  
 RT domain-containing proteins."  
 RL Mol. Cell. Neurosci. 11:149-160 (1998).  
 RN [6]  
 RN SEQUENCE OF 463-470; 503-510; 514-526; 644-665 AND 875-881,  
 RP INTERACTION WITH LMP2A, AND MUTAGENESIS OF CYS-871.  
 RC TISSUE=B-cell;  
 RX MEDLINE=20501262; PubMed=11046148;  
 RA Winberg G., Matekova L., Chen F., Plant P., Rotin D., Gish G.,  
 RA Ingham R., Erbenberg I., Pauson T.,  
 RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3  
 RT protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases."  
 RL Mol. Cell. Biol. 20:8526-8535 (2000).  
 RN [7]  
 RN INTERACTION WITH CBLC, AND PHOSPHORYLATION.  
 RP MEDLINE=22323223; PubMed=12226085;  
 RX Courbado J.-R., Fiore F., Adelaide J., Borg J.P., Birnbaum D.,  
 RA "Interaction between two ubiquitin-protein isopeptide ligases of  
 RT different classes, CBLC and AIP4/ITCH."  
 RL J. Biol. Chem. 277:45267-45275 (2002).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
 CC then directly transfers the ubiquitin to targeted substrates.  
 CC Regulates the transcriptional activity of several transcription  
 CC factors, and probably plays an important role in the regulation of  
 CC immune response. Downregulates Epstein-Barr virus LMP2A activity

CC in B cell signaling.  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Interacts via its WW domains with DRPLA, NFE2 and CBLC.  
 CC Interacts with Epstein-Barr virus LMP2A. Interacts with NOTCH1,  
 CC ICCLN, JUN and JUNB (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q96J02.1; Sequence=Displayed;  
 CC Note=No experimental confirmation available;  
 CC Name=2;  
 CC IsoId=Q96J02.2; Sequence=VSP\_008451;  
 CC TISSUE SPECIFICITY: Widely expressed.  
 CC -!- PTM: Phosphorylated on tyrosine residues.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -!- SIMILARITY: Contains 4 WW domains.  
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to erroneous  
 CC gene model prediction.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 CC EMBL; AF095745; AAK393399.1; -;  
 CC EMBL; AB066663; BAB39389.1; -;  
 CC EMBL; AL109923; CAC09387.2; -;  
 CC EMBL; AL356299; CAC09530.1; ALT\_SEQ.  
 CC EMBL; BC006848; AAH06848.1; -;  
 CC EMBL; BC011571; AAH11571.1; -;  
 CC EMBL; AF038564; AAC04845.1; -;  
 CC Genes; HGNC:13890; ITCH.  
 CC MIM; 606409; -;  
 CC InterPro; IPR000569; HECT\_domain.  
 CC Pfam; PF00632; HECT; 1.  
 CC PROSITE; PS00499; C2 DOMAIN 1; FALSE\_NEG.  
 CC PROSITE; PS00004; C2 DOMAIN 2; 1.  
 CC PROSITE; PS0237; HECT; 1.  
 CC PROSITE; PS01159; WW DOMAIN 1; 4.  
 CC PROSITE; PS00020; WW DOMAIN 2; 4.  
 CC Ubl conjugation pathway; Ligase; Nuclear protein; Repeat;  
 KW Phosphorylation; Alternative splicing.  
 FT DOMAIN 5 99  
 FT C2 DOMAIN.  
 FT DOMAIN 326 359  
 FT WW 1.  
 FT DOMAIN 359 391  
 FT WW 2.  
 FT DOMAIN 438 471  
 FT WW 3.  
 FT DOMAIN 478 511  
 FT WW 4.  
 FT DOMAIN 589 603  
 FT HECT.  
 FT DOMAIN 259 264  
 FT POLY-PRO.  
 FT BINDING 871 871  
 FT UBIQUITIN (BY SIMILARITY).  
 FT VARSPLIC 159 200  
 FT NGVSLCLPRLCNSAISAHCNLCPLGLSDSPISASRVAGFT  
 FT G -> S (in isoform 2).  
 FT /FTId=VSP\_008451.  
 FT C->A; LOSS OF UBIQUITIN PROTEIN LIGASE  
 FT ACTIVITY.  
 FT MUTAGEN 871 871  
 FT C->A; 677A2043C7B67BC CRC64;  
 FT SQ  
 FT SEQUENCE 903 AA; 102802 MW; 677A2043C7B67BC CRC64;  
 Query Match 33.2%; Score 1288.5; DB 1; Length 903;  
 Best Local Similarity 33.8%; Pred. No. 2.9e-86;  
 Matches 312; Conservative 128; Mismatches 239; Indels 245; Gaps 29;  
 QY 6 KRLTLVCLNK-ILAKDSEFRDPFPAKIVDVGSGGCHSTDTVNTLDPKWNQHYDLYVGK 64  
 DB 19 QLQITVISAKLXENKKNMFG-PSPIYEVTVTD--GOSKTEKCNNTNSPAKQPLTVIYTP 75  
 QY 65 TDSITISVNHKKIHK-----KQAGFLG-- 88

```
Db 76 VSKLHFRVSWHSLKSDVLLGTAALDIYETLKSNNMKLEEVVVTLQLGDKDEPTETIGDL 135
QY 89 --CV-----RLSNALSRKDTGY-----RLDL-----CKL-----NPSDSDAV 121
Db 136 SICLDGLQLESEWINGETTCSENGVSLCLPLECNSAISAHNCNCLPLGLSDSPISASRV 195
QY 122 RQIVUS-----LQTFD--RITGGS-----VVDCEGL-----LENGTGVYEDSGP 160
Db 196 AGFTGASQNDGSRKDETRVSTNGSDPDAGAGENRVSGNNPSPISNGG--FKSPRP 253
QY 161 GRPLSCFMBEAPFYDSTGAAGGNCRFVES-----DQGRVYVYDHFVEKRTTWDRPE-- 358
Db 254 PRP-----SRPPPTPRPASVNGSPSATSESDGSGTSLPPTNTNTSEGATSGLLIP 308
QY 193 -----PSQORLOAQLRNPVGRSL-----QTPQNRPHCH 223
Db 309 LTISSGSGRPLNPVQAPLPGEWQRV-----DQGRVYVYDHFVEKRTTWDRPE-- 358
QY 224 QSPPELPEGEQRTTQGGQVYFHTQTGSTWHDPRIP-----PSEGSLEDEELPAQRYERD 343
Db 471 -SQQLNE--KPLPGWEMRFTVDGIPYVDHNRRITTYIDRTGKSALDNGPQIAYVRD 527
QY 344 LVOKLKLVR---HELSLOQPQAGHCRIEVSREIFEESYRQIMQVRPKDLKRLMKVFRG 400
Db 528 FXAKVQYFRFWCQQLAMPO---HIKITVTRKTLFEDSFQQIMSPSQDRLRRLWVIFPG 583
QY 401 EBGLOYGVARWELVLCHEMLNPXYGLFOYSTDNIYMLQINPDSINPDHLSYFHFYGR 460
Db 584 EBGLOYGVARWELVLSHEVLNPXYGLFOYAGKDNVCLQINPASYINPDHLYFRFTR 643
QY 461 INGLAVFHGYINGGTVPFYKQLCKPTQLSDLESVDPELHSLVLTLENDITPV-LDH 519
Db 644 FIAMALFHGKIDTGFSPFYKRLNKPVLGKDLJESIDPEFVNSLIWKENNIECDLEM 703
QY 520 TFCVEHNAFRLQHLKPNRNPVTEENKEYVRLVYNRPFMGIEAQFLAQKNGNE 579
Db 704 YFSVDKEILGETKSHDLKNGCNILVTEENKEYIRMAENRSLSGVEEQTAFFEGFNE 763
QY 580 LIPQHLKPFDOKELELIIGGLDKIDLNDKSNTRLKHCVADSNIVRWFQAVETFDER 639
Db 764 ILPQVLYQYFDAKEVLLCGQEDLDNDQWQHAIYRYARTSKQIMFWQFVKEIDNEK 823
QY 640 RARLLQFVTGSTRVPLQGFALQSGTGAAPRLFTIHLIDANTDNLKPAHTCFNRIDIPP 699
Db 824 RMELLQFVTGTRCLPVGGFADLMGNS--GPOKFCIEKV-GKENWLPRTSHTCFNRDLDP 879
QY 700 YESYKLYEKLTAVEETCGRAVE 723
Db 880 YKSYQLKEKLTAETEETGFGQE 903

RESULT 9
ID NED4 HUMAN STANDARD; PRT; 1000 AA.
AC P46334;
DT 01-NOV-1995 (Rel. 32, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).
GN NEDD4 OR KIAA0093.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP
RA Hinz U.;
RL Unpublished observations (JUN-2003).
RN [2]
RC SEQUENCE OF 74-1000 FROM N.A.
RX MEDLINE=95308325; PubMed=7788527;
RA Nagase T., Miyajima N., Tanaka A., Suzuki T., Seki N., Sato S.,
RA Tabata S., Ishikawa K.-I., Kawabayashi Y., Kotani H., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. III.
RT The coding sequences of 40 new genes (KIAA0091-KIAA0120) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
EL DNA Res. 2:37-43(1995).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
CC then directly transfers the ubiquitin to targeted substrates (By
CC similarity).
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Interacts with UBE2D2. Binds SCNN1A, SCNN1B and SCNN1G
CC (By similarity).
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thiolester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 4 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -!- CAUTION: The sequence of the N-terminus was deduced from the
CC genomic sequence.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AC039057; -; NOT ANNOTATED_CDS.
CC EMBL; D42055; BAA07655.1; -;
CC HSSP; Q13526; 1PIN.
CC Genew; HGNC:7727; NEDD4.
CC MIN; 602278; -.
CC InterPro; IPR000008; C2.
CC InterPro; IPR008973; C2 CaLB.
CC InterPro; IPR000569; HECT domain.
CC InterPro; IPR001202; WW_Rap5_WMP.
CC Pfam; PF00168; C2; 1.
CC Pfam; PF00632; HECT; 1.
CC Pfam; PF00397; WW; 4.
CC SMART; SM00239; C2; 1.
CC SMART; SM00119; HECTC; 1.
CC SMART; SM00456; WW; 4.
CC PROSITE; PS00499; C2_DOMAIN_1; 1.
CC PROSITE; PS00004; C2_DOMAIN_2; 1.
CC PROSITE; PS00237; HECT; 1.
CC PROSITE; PS01159; WW_DOMAIN_1; 4.
CC PROSITE; PS00020; WW_DOMAIN_2; 4.
CC Ubl conjugation pathway; Ligase; Repeat.
CC DOMAIN 9 44
CC FT DOMAIN 77 81
CC FT DOMAIN 106 207
CC FT DOMAIN 285 288
CC FT DOMAIN 291 324
CC FT DOMAIN 448 481
CC FT DOMAIN 521 554
CC FT DOMAIN 573 606
CC FT DOMAIN 665 1000
CC FT DOMAIN 967 967
CC FT BINDING
CC SEQUENCE 1000 AA; 114936 MW; 372808850C149CB CRC64;
CC
CC Query Match 32.4%; Score 1259; DB 1; Length 1000;
CC Best Local Similarity 33.6%; Pred. No. 4.9e-84;
```

Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;

```

QY      7 IRLTVLCAKNLAKXDFRLPDPFAKIV---VDGSGQCHSTDIVKNTLDPKWN-----55
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     121 VVRVVIAGIGLAKKDILGASDPVVRVTLVDPMNGVLTSVOTKTIKSLNPKWNEILLPRV 180
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      56 ---QH-----YDLVYKTD--SITISVN-----HKIHKQOG 83
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     181 HPQQRHLLEFVFDENRLTRDFFLGQVDVPLYPPLPTENPELERPYTFKDFVLHPSHKSERV 240
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      84 AGFLGCVRLLSNAISRLLKDT-----GYQRLD-----LCKL-----113
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     241 KGYL-----RLKMTYLPKTSGEDDAAEAEEPCWVVLVDQDAACHLQQQQE 289
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      114 -----NPSD--TDVAVRQIVVSILQ-----TRDR 134
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     290 PSPPLPGWEERQDILGRITVYNHESRRTQWKRPDPQDNLTDAENGNIQLQAFRTTTRRQ 349
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      135 IGTGSSVDDCGILEN-----EGTVYEDSG-PGRPLSCFMEEPAPYTD-----176
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     350 ISEETESVDNOESSENWEIITREDAETWYSQAFSPFPSNLDVTHLAEELNARLITFG 409
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      177 ---STGAAAGGNCR-----FVSPS-----QDRLQAQRL-- 204
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     410 NSAVSQPASSNHSSRGSLQAYTFEEOPTPLVLLPTSSGLPPGWEEKQDGRGRSYVDH 469
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      205 -----RNPVDRGSLQTPQ-----NPHGHQSPPELPEGYEQ 234
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     470 NSRTTWTWKTPVQATVSTLSQTSQSAGSQSASTSDSGQVTPSPSEIOGFLPKGMEV 529
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      235 RTTVQGVYFLHTQTGVSTVHWDPR--IPRDLNSV-----NCDLGLPLPGWEVRSTVSGRI 288
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     530 RHAPNGRPFIDHNTKTTTWEDPELKI PAHLRGKTSLSLTDNDLGPLPGWEERTHTDGR 589
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      289 YFVDHNRTTQFTDPRLHHIMHQCQLKPSQPLPLPSGSLDEBELPAQYERDLVQKL 348
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     590 FYINHNKRTQWEDPRELV-----AITGPAPV-----YSRDYRKKY 626
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      349 KVLRHLSLQOPQACHRIEVSREETFEESYROIIMKMRPKD-LKKRLMVKFRGEEGLDYG 407
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     627 EFRFRKLKKNDIENKPEMKLRATVLEDSYRIMGVRAADFLKARLWTFDGEKGLDYG 686
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      408 GVAREWLYLCHEMLNPYIGLFQYSTDNIVMLQINPDSSI-NPDHLSYFHFVGRIMGIAV 466
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     687 GVAREWFLISKENFNFPYIGLFYSATDNYTLQINPNSGLCNEDHLSYFKFTIGRVAGMAV 746
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      467 PHGYVINGCFVVPYKOLLGKPIQLSDLESVDPELHKSILVWILENDITFVLDDHFCVEHNN 526
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     747 YHGKLLDGGFFIRPYKMWLHKPI TLHDMSVDSVEYNSLRWILENDPTE-LDLRFIDE 805
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      527 AFGRILQHELKPNRNPVPTVEENKKEYVRLVYNRMFRMGIEAQFLALQKGFNELIPQHL 586
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     806 LFGQTHQHELKNGSGEIVWTNKNKKEYIYLIQWRFNRIQKMAAFKEGFPELLIPQDL 865
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      587 KPFDOKELELLIGLDKLDNDKWSNTRLKH-CVADSNIVRWFQAVETPDEERPARLLO 645
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     866 KIFDNELELMLCGLDVNDNREHTKYKNGYSANHQVIQWPKAVLMMDSKRIURLQ 925
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      646 FVTGSTRVPLQGFKAALOGSTGAAGPRFTTHLDANTDLNPKAHTCFNRIDIPPYESYEK 705
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     926 FVTGSTRVPMNGFAELYGSN---GPQSFTVEQW-GTPEKLPRAHCTFNLDLPIPPYSEEE 981
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY      706 LYEKLLTAVETQGF 720
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db     982 LWDKLQMAIENTQGF 996
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RX MEDLINE=22338357; PubMed=12450395;
RA Galinier R., Gout E., Iortat-Jacob H., Wood J., Chroboczek J.;
RT "Adenovirus protein involved in virus internalization recruits
RL Biochemistry 41:14299-14305(2002).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 546-917, MUTAGENESIS OF
RP GLU-614; HIS-621; ASP-675; GLU-798; MET-804; ARG-845; GLN-848
RP AND ARG-855 AND FUNCTION.
RX MEDLINE=22423789; PubMed=12535537;
RA Verdecia M.A., Jazeiro C.A.P., Wells N.J., Ferrer J.-L., Bowman M.E.,
RA Hunter T., Noel J.P.;
RT "Conformational flexibility underlies ubiquitin ligation mediated by
RT the WPI1 HECT domain E3 ligase."
RL Mol. Cell 11:249-259(2003).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
CC then directly transfers the ubiquitin to targeted substrates.
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Binds KLF2 (By similarity). Binds SCNN1A, SCNN1B, SCNN1G,
CC WBP1, WBP2, DRP1A and adenovirus type 2 PIII.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=A;
CC IsoId=Q9H0M0-1; Sequence=VSP\_007601, VSP\_007603;
CC Name=2; Synonyms=B;
CC IsoId=Q9H0M0-2; Sequence=VSP\_007602;
CC Name=3; Synonyms=C;
CC IsoId=Q9H0M0-3; Sequence=VSP\_007602;
CC Name=4; Synonyms=D;
CC IsoId=Q9H0M0-4; Sequence=Not described;
CC Name=5; Synonyms=E;
CC IsoId=Q9H0M0-5; Sequence=Not described;
CC Name=6; Synonyms=F;
CC IsoId=Q9H0M0-6; Sequence=VSP\_007600;
CC -!- TISSUE SPECIFICITY: Detected in heart, placenta, pancreas, kidney,
CC liver, skeletal muscle, bone marrow, fetal brain, and at much
CC lower levels in adult brain and lung. Isoforms 1 and 5 predominate
CC in all tissues tested, except in testis and bone marrow, where
CC isoform 5 is expressed at much higher levels than isoform 1.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thiolester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 4 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).
CC
CC -----
CC EMBL; A1136739; CAB66673.1; -
CC EMBL; AY043361; AAK94668.1; -
CC EMBL; BC015380; AAH15380.1; -
CC EMBL; BC036065; AAH36065.1; -
CC EMBL; U96113; AAC51324.1; -
CC PDB; 1ND7; 23-SEP-03.
CC MIM; 602307; -
CC GO; GO:000151; C:ubiquitin ligase complex; NAS.
CC GO; GO:0003515; F:protein binding; IPI.
CC GO; GO:0004842; F:ubiquitin-protein ligase activity; TAS.
CC GO; GO:0007417; P:central nervous system development; NAS.
CC GO; GO:0030324; P:lung development; ISS.
CC GO; GO:0016481; P:negative regulation of transcription; ISS.
CC GO; GO:0016567; P:protein ubiquitination; TAS.
CC GO; GO:0007165; P:signal transduction; NAS.
CC GO; GO:0030217; P:cell differentiation; ISS.
CC GO; GO:0046718; P:viral entry; TAS.

DR InterPro: IPR000008; C2.
DR InterPro: IPR008973; C2\_CalB.
DR InterPro: IPR000569; HECT\_domain.
DR InterPro: IPR002349; WW.
DR InterPro: IPR001202; WW\_Rsp5\_WWP.
DR Pfam: PF00168; C2; 1.
DR Pfam: PF00632; HECT; 1.
DR Pfam: PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS00499; C2\_DOMAIN\_1; FALSE\_NEG.
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.
DR PROSITE; PS00237; HECT; 1.
DR PROSITE; PS01159; WW\_DOMAIN\_1; 4.
DR PROSITE; PS00020; WW\_DOMAIN\_2; 2.
KW Ubl conjugation pathway; Ligase; Repeat; Alternative splicing;
KW 3D-structure.
FT DOMAIN 5 98 C2 DOMAIN.
FT DOMAIN 349 382 WW 1.
FT DOMAIN 381 414 WW 2.
FT DOMAIN 456 489 WW 3.
FT DOMAIN 496 529 WW 4.
FT DOMAIN 588 922 HECT.
FT BINDING 890 890 UBIQUITIN (BY SIMILARITY).
FT VARSPPLIC 23 240 Missing (in isoform 6).
FT VARSPPLIC 112 120 LERVKQLK -> CWLKKARME (in isoform 2).
FT VARSPPLIC 112 242 LERVKQLESLNKKNGIAQTGLTVVLDGLVIEQENITNC
FT VARSPPLIC 121 922 LERVKQLESLNKKNGIAQTGLTVVLDGLVIEQENITNC
FT MUTAGEN 614 614 E-A: REDUCES UBIQUITIN TRANSFER.
FT MUTAGEN 621 621 H-A: STRONGLY REDUCES UBIQUITIN
FT MUTAGEN 675 675 D-A: REDUCES UBIQUITIN TRANSFER.
FT MUTAGEN 798 798 E-A: REDUCES UBIQUITIN TRANSFER;
FT MUTAGEN 804 804 M-P: STRONGLY REDUCES UBIQUITIN
FT MUTAGEN 806 806 TRANSFER; WHEN ASSOCIATED WITH P-806.
FT MUTAGEN 845 845 E-S: STRONGLY REDUCES UBIQUITIN
FT MUTAGEN 848 848 TRANSFER; WHEN ASSOCIATED WITH P-804.
FT MUTAGEN 848 848 Q-A: ABOLISHES UBIQUITIN TRANSFER; WHEN
Query Match 31.9%; Score 1239; DB 1; Length 922;
Best Local Similarity 32.5%; Pred.No.1.3e-92;
Matches 300; Conservative 138; Mismatches 282; Indels 224; Gaps 25;
QY 5 IKRLTVLCARNAKQDFRLPDPFAKIVVDGSGQCHSDTVK--NTLDPKRNQHYDLYV 62
DB 18 LQLQVTSSAKLRKKNWFGTA-IYTEVVDVG---BITKTAKSSSSSNPKWDQLTVNV 72
QY 63 GKTDSTIISVWNHKKIHKQAGELG-----CVLELNSAISRKLD----- 102
DB 73 TPQTTLEFQVWSHRTL---KADALLGKATIDLKALLIHRKLERVKEQKLSLENKNGI 129
QY 103 --TG-----YQRDLCKINPS-----DTDAVRGQIVVSLQTRDRI---GTGG- 139
DB 130 AQTGLTIVLDGLVIEQENITNCSSPTIEIQENGDAHENGESPSARTTARLAVEGTNGI 189
QY 140 -SVVDCRGLLENEGTVVEDSGRPLSCFM-----EEPAP-----Y 174
DB 190 DNHVPTSTLVQNSCCSYVNGDNTFSSFSQVAARPKNTAPKPLASEPADDTVNGESSSF 249
QY 175 TDSGTAAAGG-----GNC-----RFVESPSQDORLOAQR 203

Db 250 APTDNASVTGTPVSEENALSPNCTSTTVEDPPVQEIILTSSENNECIPSTSAEISEARS 309  
 QY 204 LRENDVRGSLT-----PQNRPHGHOSP-----ELREGYBORTTVOGVYFLH 246  
 Db 310 ILEPDTNSRRSSAFEAARKSQPCGMDPVRRQSGNANTETLPSGWQRKDPGRRTYYVD 369  
 QY 247 TQTGVSTWHDPR-IP-----260  
 Db 370 HNRITTTWERQPLPGWERRVDRRRVYVDHNRITTTTQCRPTSVRNFQWQSRNQ 429  
 QY 261 -----RDLSVNC-----DELGLPGWEVRSTVGRIFYVDHNRRTTQFTDPR 305  
 Db 430 LQAMQOQNORYLYSASMLAENDPYGLPPGWEKRVDSITDRVYFNHNTKTQWEDPRT 489  
 QY 306 HHIMNHQ-----CQLKPSQPLPLPSBGSLEDEELPAQRY 340  
 Db 490 QGLQNEEPLPGWEIRYTRGVRVYVDHNTTTTFKDPNGKSVTKGG-----PQIAY 543  
 QY 341 ERDLVQKLVLHLSLQOQAGHCRIEVSEELFESSYQIMQMRPDKLKLMLVKFRG 400  
 Db 544 ERGFELKLAHPRY-LCOSNALPSHVKNVSRQTLFDSFQIIVALKPYDLRRRLYVIFRG 602  
 QY 401 BEGLDYGVAEWELLYLICHEMLNPYGLFOYSTDNIMYLOINPDSSINPDHLSYEFHVR 460  
 Db 603 BEGLDYGGLAREWFLLSHEVLNPMYCLFVYAGKNYCLQINPASTINPDHLSYFCFGR 662  
 QY 461 IMGLAVFHGVIYNGFTVPVYKQLLKPQIOLSDLESVDPELHLSVLWILENDITPV-LDH 519  
 Db 663 FIAMALFHGKFDITGFSLPYKRLSKLTKDLESIDTFYNSLIWIRONNTEECGLEM 722  
 QY 520 TFCYEHNAFCHLOHELKPNQNVPTTEENKKEVRLVYVNRWFRGTEAFQALQGFNE 579  
 Db 723 YFSVDMELKGVTHDGLKSGNSTLVTEENKDEYIGLMTWERSRGVQEQTKAPLQGFNE 782  
 QY 580 LIPQHLKPFQDKLELIIGGLDKIDLNDWKSNTLKHCVADSNIVRWFWQAVETDEER 639  
 Db 783 VVPLQWLQYFDEKELEWVLCMQEVDLADQCRNTVYHYHRTNSKQIINFQVKETDNEV 842  
 QY 640 RARLQVFTGSTRVPLQGFALQGSTGAAGRPFTIHLIDANTDNLPAKTCNRLDIPP 699  
 Db 843 RMRLQVFTGSTRVPLQGFALQGSTGAAGRPFTIHLIDANTDNLPAKTCNRLDIPP 699  
 QY 700 YESYEKLVEKLLTAVETTCQFAVE 723  
 Db 899 KYSEQLKELLFAIETEGFQE 922

## RESULT 11

## WBP2 MOUSE

AC WBP2 MOUSE STANDARD; PRT; 870 AA.  
 ID Q9BHO; O8BTG4; Q923F6;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Nedd-4-like ubiquitin-protein ligase WBP2 (EC 6.3.2.-) (WW domain-  
 containing protein 2).  
 GN WBP2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OK NCBI\_TaxID=10090;  
 RN  
 RP  
 RX MEDLINE=22354683; PubMed=12456851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Osato N., Saico R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Butt C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,  
 RA Kongaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maitais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Nunata K., Okido T., Pavan W.J., Perte G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilning L.G., Wynshaw-Boris A., Yangisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Korn H., Nakamura M., Sakazume N., Sato K.,  
 RA Shitaki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs";  
 RL Nature 420:563-573(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6, and FVB/N; TISSUE=Brain, and Breast tumor;  
 RX MEDLINE=22386257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Frange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts  
 CC ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a  
 CC thioester and then directly transfers the ubiquitin to targeted  
 CC substrates (By similarity).  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Binds SCNN1A, SCNN1B, SCNN1G, WBP1, WBP2 and DEPLA (By  
 CC similarity).  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thioester formation.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 4 WW domains.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AK004962; BAB23702.1; -  
 CC EMBL; AK088936; BAC40661.1; -  
 CC EMBL; AK090392; BAC41195.1; -  
 CC EMBL; BC004712; AAH04712.1; -  
 CC EMBL; BC039921; AAH39921.1; -  
 CC EMBL; BC048184; AAH48184.1; -  
 CC



RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyx S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Small D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RA "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP INTERACTION WITH KLF2.  
RX MEDLINE=2336920; PubMed=11375995;  
RA Konkrig M.D., Wani M.A., Lingrel J.B.;  
RT "Lung Kruppel-like factor contains an autoinhibitory domain that  
RT regulates its transcriptional activation by binding Wp1, an E3  
RT ubiquitin ligase.";  
RL J. Biol. Chem. 276:29299-29306(2001).  
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
CC then directly transfers the ubiquitin to targeted substrates (by  
CC similarity).  
CC -!- PATHWAY: Ubiquitin conjugation; third step.  
CC -!- SUBUNIT: Binds SCNN1B, SCNN1G, WBP1, WBP2, DRPLA and  
CC adenovirus type 2 PIII (By similarity). Binds KLF2.  
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
CC thiolester formation (by similarity).  
CC -!- SIMILARITY: Contains 1 C2 domain.  
CC -!- SIMILARITY: Contains 4 WW domains.  
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
CC domain.  
CC -!- CAUTION: Ref.1 (BAC38473) sequence differs from that shown after  
CC position 391; the second part of that clone maps to another  
CC chromosome.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AK033138; BAC28168.1;  
CC EMBL; AK082346; BAC38473.1; ALT SEQ.  
CC EMBL; BC021470; AAH21470.1; ALT\_INIT.  
CC EMBL; BC051405; AAH51405.1;  
CC MGD; MGI:1861728; Wp1.  
CC GO; GO:0000151; C:ubiquitin ligase complex; ISS.  
CC GO; GO:0005515; F:protein binding; IPI.  
CC GO; GO:0004842; F:ubiquitin-protein ligase activity; ISS.  
CC GO; GO:0007417; P:central nervous system development; ISS.  
CC GO; GO:0030324; P:lung development; TAS.  
CC GO; GO:0016481; P:negative regulation of transcription; IDA.  
CC GO; GO:0016587; P:protein ubiquitination; ISS.  
CC GO; GO:0007165; P:signal transduction; ISS.  
CC GO; GO:0030217; P:T-cell differentiation; TAS.  
CC GO; GO:0046718; P:viral entry; ISS.  
CC InterPro; IPR000008; C2.  
CC InterPro; IPR008973; C2 CALB.  
CC InterPro; IPR000569; HECT\_domain.  
CC InterPro; IPR002349; WW.  
CC InterPro; IPR001202; WW\_Rsp5\_WWP.  
CC Pfam; PF00632; HECT.1.

DR Pfam; PF00397; WW; 4.  
DR PRINTS; PR00403; WWDOMAIN.  
DR SMART; SM00119; HECTc; 1.  
DR SMART; SM00456; WW; 4.  
DR PROSITE; PS00499; C2 DOMAIN 1; FALSE\_NEG.  
DR PROSITE; PS00004; C2 DOMAIN 2; 1.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW\_DOMAIN 1; 4.  
DR PROSITE; PS00020; WW\_DOMAIN 2; 4.  
DR Ub1 conjugation pathway; Ligase; Repeat.  
FT DOMAIN 5 98  
FT C2 DOMAIN  
FT DOMAIN 345 378 WW 1.  
FT DOMAIN 377 410 WW 2.  
FT DOMAIN 452 485 WW 3.  
FT DOMAIN 452 525 WW 4.  
FT DOMAIN 584 918  
FT HECT.  
FT BINDING 886 886 UBIQUITIN (BY SIMILARITY).  
FT CONFLICT 302 302 V -> L (IN REF. 1; BAC28168).  
SQ SEQUENCE 918 AA; 104593 MW; 01478A3C1CFDAA9 CRC64;  
Query Match 31.7%; Score 1232; DB 1; Length 918;  
Best Local Similarity 32.3%; Pred. No. 4.1e-82;  
Matches 302; Conservative 137; Mismatches 247; Indels 248; Gaps 25;  
QY 5 IKIRLVLCANLAKKOPFRLLDPFAKIVDGSQCHSTDTVKNLDPKMNQHYDLVYVK 64  
DB 18 LQLKVTSSAKLKRKNWFGTA-IYTEIVD--GEVKKAKSSSSSNPKMDEQLIVNVTP 74  
QY 65 TDSITISWNHKKHKKQAGFLGCVRLLSNAISRLKD---TGVRDLCLKLNPSDTEAV 121  
DB 75 QTLEFRVWSHHTL--KADA-----DLGKATVDLKVLLTHNRKL-----EKV 115  
QY 122 RQIVVSLQTRDRI--GTGGSVVDCRGLL-----EN----- 150  
DB 116 KEQLKSLNKNKNGIVQIGELIVLDGLVIEQEPVNRSSSPPIIQNGDHALHNGDPAT 175  
QY 151 -----EGTVED-----SGP-----GRPLSCFMEEP--- 171  
DB 176 RTTPRLPVEGTIGIDNHVSTNTVVPNSCCSHVANGENTPSPSQVAARPKAPKPVTS 235  
QY 172 APYDTS----- 177  
DB 236 APTSTVNGESSVLADNTSTMGILLPSDDTSTSNCTSTTQBPVQEPSESESECI 295  
QY 178 -TGAAGGNCRFVESPQDQRLQA--QRLNRP-----VSGSLQT-----PQN 218  
DB 296 PSASAEVGPARSLLTDPDSRNNSVFDPKVRQEGCVPLRPSQSGNTTEALPSCGQPK 355  
QY 219 RPHGH-----QSPV-LPEGVGEORTTVQGVYFLHTQTVSTWHPRIPRDLN 264  
DB 356 DPHGTYTVVDHNTTRTTWRFQPLPGWERRVDGRGVYVVDHNTTRTTWRFQPLPGW 415  
QY 265 -----SYNCDLGLPFGWVRSTVSGRIYFVVDHNN 295  
DB 416 FEWQSQORNQLOQAMQPNQRYLYSASMLAAENDPYGLPLPGWVKRVDSTDRVYFVNHNT 475  
QY 296 RTTQFTDRLHIMHQ-----COLKEPSQPLPSPGSL 330  
DB 476 KTTQWEDPTQGLPNEEPPEGWIRYREGVRVVDHNTTRTTTQDPKRNKSVTKGG- 534  
QY 331 EDEELPAQRYERDLVQKLKVLRLHLSLQQPOAGHCRIEVSREEFESYQIMKMRPKDL 390  
DB 535 -----PQIAYERSFRWKLAFERY-LCCSNALPSHVKNVSRQTLFEDSFQIMALKPYDL 588  
QY 391 KKRLMVKPRGEGLDYGVAVREWYLLICHEMLNPIYGLFQVSTDNITVLMQINPDSSINPD 450  
DB 589 RRLYVIFRGEGLDYGLAREWFFLLSHENVPMYCLFAYAGKNYCLQINPASTINPD 648  
QY 451 HSLVPHFVGRIMGLAVPHGHVINGFTVPFYKQLLKGKPIQLSDLESVDPELHLSLVWILE 510  
DB 649 HSLVFCFGRFIALMALFHGKFDITGFSLPFYKMLSKKLIKLESIDTFEYNSLIWIRD 708  
QY 511 NDIETV-LDHTFCVEHNAGFRILQHELKPNGRNVPTVEENKKEVRLYVNVWRVNRGIEAQ 569

```

Db 709 NNIEECLEWYSVDMEIIGKVTSHDLKUGSNILVTENKDEYIGLMTWFSRGVQEQ 768
QY 570 FLALQKGFNELIPQHLKPFDOKELELIIGGLDKIDLNDKWKSNTRLKHCVADSNIVRWFW 629
Db 769 TRAFLDGFNEVFLQWLOVDFEKELEVMICGQEVLDLADQRTVYRHYTRNSKQIWNF 828
QY 630 QAVETFDERRARLLQFVTGSTRVPLQGLKALQSGTGAAGPRLFTIHLIDANTDNLPAH 689
Db 829 QFVKETDNRVRLQFVTGSTRCLPLGGFAELMGSN--GPKQFCIEKVGKDT-WLPRSH 884
QY 690 TCFNRIDIPPSYEKLYEKLITAVEETCGFAVE 723
Db 885 TCFNRDLDPYKSYEQLEKLLFAITEETGFGQE 918

RESULT 13
NED4 RAT
ID NED4 RAT STANDARD; PRT; 887 AA.
AC Q62940;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-).
GN NEDB4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A., AND INTERACTION WITH SCNN1A; SCNN1B AND SCNN1G.
RC TISSUE=Lung;
RX MEDLINE=96221297; PubMed=8665844;
RA Staub O., Dho S., Henry P., Correa J., Ishikawa T., McGlade J.,
RA Rotin D.;
RT "WW domains of Nedd4 bind to the proline-rich PY motifs in the
RT epithelial Na+ channel deleted in Liddle's syndrome.";
RL EMBO J. 15:2371-2380(1996).
RN [2]
STRUCTURE BY NMR OF 452-499 IN COMPLEX WITH SCNN1B, AND INTERACTION
RP WITH SCNN1A; SCNN1B AND SCNN1G.
RX MEDLINE=2123577; PubMed=11323714;
RA Kanellis V., Rotin D., Forman-Kay J.D.;
RT "Solution structure of a Nedd4 WW domain-ENAC peptide complex.";
RL Nat. Struct. Biol. 8:407-412(2001).
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and
CC then directly transfers the ubiquitin to targeted substrates (By
CC similarity).
CC -!- PATHWAY: Ubiquitin conjugation; third step.
CC -!- SUBUNIT: Interacts with UB52D2 (By similarity). Binds SCNN1A,
CC SCNN1B and SCNN1G.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Ubiquitous. Expression is highest in lung,
CC kidney and brain.
CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
CC thioester formation.
CC -!- SIMILARITY: Contains 1 C2 domain.
CC -!- SIMILARITY: Contains 3 WW domains.
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase
CC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U50842; AAB48949.1; -
CC PIR; S70642; S70642.
CC PDB; 115H; O2-MAY-01.

```

```

DR InterPro; IPR000008; C2.
DR InterPro; IPR00973; C2 CalB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 3.
DR PRINTS; PR00360; C2DOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS50020; WW_DOMAIN_2; 3.
KW Ub1 conjugation pathway; Ligase; Repeat; 3D-structure.
FT DOMAIN 62 163 C2 DOMAIN.
FT DOMAIN 246 279 WW 1.
FT DOMAIN 402 435 WW 2.
FT DOMAIN 459 492 WW 3.
FT DOMAIN 551 887 HECT.
FT BINDING 854 854 UBIQUITIN (BY SIMILARITY).
SQ SEQUENCE 887 AA; 102394 MW; D74B1097688CD9A1 CRC64;

Query Match 31.6%; Score 1227.5; DB 1; Length 887;
Best Local Similarity 33.9%; Pred.No. 8.3e-82;
Matches 296; Conservative 127; Mismatches 224; Indels 227; Gaps 28;

QY 7 IRLTVLCAKNAKAKFFRLPDPFAKIV-----VPGSGCHSTDTVKNTLDPKWN----- 55
Db 77 VRVKVIAGIGLAKKIDLGASDPYVRVTLDPMSGLTSTVQTKTKKSLNPKWNEILFRV 136
QY 56 ---QH-----YDLYVGKTD--SITISVWN-----HKIHKKQG 83
Db 137 LPQHRILFEVDENRLTRDQGLQVDVPLPTFENRPERPYTFKDFVLHPRSHKSRV 196
QY 84 AGFLCGVLLLSNAISRKLT-----RLKMTYLPKNGSDDENADQASELEFGWVLDQPDAAHLQHPPEP 245
Db 197 KGYL-----GY-QRLDLC-----KLNPSD--TDVVGQIVV-----SLQTRDRI 135
QY 104 -----GY-QRLDLC-----KLNPSD--TDVVGQIVV-----SLQTRDRI 135
Db 246 SPLPFGWEERQDVLGRTYYVNHESRTTQWKRPSPEDLTDENGDIQLAHGAFTRRQI 305
QY 136 GTGSGVVDRCG-----LLENEGTVYEDSGGRPLS-----C 166
Db 306 SE-----DVGDPDNHESPENWEIVREDNTIYSGAVQSPSPGHPDVQVZLASELOTRIT 360
QY 167 FMEPPA---PYTDSTGAACGGN--CRFVESPQDQRLQARLNPDPVRSLOTPQNRPH 221
Db 361 MYGNPATSQPVTSNHSRSGSSQTCIFERQPTLPVLLPT-----RIPRDLNS---VNCDE 270
QY 222 GHQSPELPEGYEQRTTVQGVYFLHTQTGVSTVHDP-----RIPRDLNS---VNCDE 270
Db 401 ---SSGLPFGWEERQDQGRSRSYVDHNSKTTTWSKPTMQDDPRSKIPAHURGTVPVDSND 457
QY 271 LGPLPPGWEVRSTVSGRIYFVDHNNRTQTDFRLHHIMNHQCLKEPSQFLPLPSRGL 330
Db 458 LGPLPPGWEERTHTDGRVFFINHNKKTOWEDPRMQNV-----AITGPAEP----- 503
QY 331 EDELPQARVERDLVQKLKVLRHLSLQOQAQCHIEVSRREEIFESYQIMKRPKD- 389
Db 504 -----YSRDYKRYEYFRRLKKQTDPNKKFKURRANILEDSYRIRNGVRADF 554
QY 390 LKRLMKYKFEEGELDYGGVAREWLYLLCHEMLNPPYGLFOYS--TONIYMLQINPSSI- 447
Db 555 LKARLWTEFDGKGLDYGGVAREWFFLISKEMFNPPYGLFEYATEDNYTLQINPNSGLC 614
QY 448 NPDHLSYTFHVRIMGLAVPHGYINGGFTVPYKQLLGPILQSLDLESVDPELHSLVW 507
Db 615 NEDHLSYFKFIRVAGNAVYHGKLLDGGFFIRFPYKWLQKLITLHDWESVDSEYSSLRW 674

```

QY 508 ILENDITFVLDHTFCVHNAGRILOHELKPNGRNVPVTENKKEYVRLVYVNRFRMGIE 567  
 DB 675 ILENDPTE-LDLRIFIDIELFGQTHLKTGSEVVVTKNKKKEYILYVIOWRFNRIQ 733  
 QY 568 AQFALOKGNEELIPQHLKPFDDOKLELLIGGLDKIDLNDKSNTRLKHCVA-DSNIVR 626  
 DB 734 KOMAAFEKGFPELLIPQDLIKFIDNELEMLCGLGDVDVNDWREHTKYKNGYSLNQVH 793  
 QY 627 WFOQAVETFDERRARLLQFVTGSTRVPLQFGKALQGSTGAAGPRFTTHLIDANTDNL 686  
 DB 794 WFKAVLMDSEKIRLLQFVTGSTRVPMNGFAELYSN---GQPSFTVQW-GTDPKLP 849  
 QY 687 KAHTCFNRIIDPPVESYKLYELLTAVESTCGF 720  
 DB 850 RAHTCFNRLDLPFVSEFDELWDLQMAIENTQGF 883

RESULT 14  
 WWP2\_HUMAN STANDARD; PRT; 870 AA.  
 AC 000308; Q96C22; Q9BWN6;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Nedd-4-like ubiquitin-protein ligase WWP2 (EC 6.3.2.-) (WW domain-  
 containing protein 2) (Atropin-1 interacting protein 2) (Aip2).  
 GN WWP2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 (1)  
 RP SEQUENCE FROM N.A., AND INTERACTION WITH WBP1; WBP2; SCNN1A; SCNN1B  
 RP AND SCNN1G.  
 RC TISSUE=Brain, and Placenta;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Heish F.,  
 RA Diatchenko M., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Heltan E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Truchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Goughwood J., Schmutz J., Myers R.M.,  
 RA Butterfield V.S.N., Kryzyski M.I., Skalska U., Smalls D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 (3)  
 RP INTERACTION WITH DRPLA, AND TISSUE SPECIFICITY.  
 RX MEDLINE=98313405; PubMed=9647693;  
 RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,  
 RA Kaminsky Z., Kleiderlein J.J. Jr., Sharp A.H., Ross C.A.,  
 RT "Atrophin-1, the DRPLA gene product, interacts with two families of  
 RT WW domain-containing proteins."  
 RL Mol. Cell. Neurosci. 11:149-160(1998).

RN [4]  
 RP INTERACTION WITH SCNN1A; SCNN1B AND SCNN1G.  
 RX MEDLINE=22157111; PubMed=12167593;  
 RA McDonald F.J., Western A.H., McNeill J.D., Thomas B.C., Olson D.R.,  
 RA Snyder P.M.;  
 RT "Ubiquitin-protein ligase WWP2 binds to and downregulates the  
 RT epithelial Na(+) channel."  
 RL Am. J. Physiol. 283:F431-F436(2002).  
 (5)  
 RN INTERACTION WITH ADENOVIRUS TYPE 2 PIII.  
 RP MEDLINE=22338357; PubMed=12450395;  
 RX Galinier R., Gout E., Lortat-Jacob H., Wood J., Chroboczek J.;  
 RA "Adenovirus protein involved in virus internalization recruits  
 RT ubiquitin-protein ligases."  
 RL Biochemistry 41:14299-14305(2002).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts  
 CC ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a  
 CC thioester and then directly transfers the ubiquitin to targeted  
 CC substrates (By similarity).  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Binds SCNN1A, SCNN1B, SCNN1G, WBP1, WBP2, DRPLA and  
 CC adenovirus type 2 PIII.  
 CC -!- TISSUE SPECIFICITY: Detected in heart, throughout the brain,  
 CC placenta, lung, liver, muscle, kidney and pancreas.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 4 WW domains.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U96114; AAC51325.1; -;  
 CC GO; GO:0000151; C:ubiquitin ligase complex; TAS.  
 CC GO; GO:0004842; F:ubiquitin-protein ligase activity; TAS.  
 CC GO; GO:0006464; P:protein modification; TAS.  
 CC InterPro; IPR000008; C2.  
 CC InterPro; IPR008973; C2\_CaLB.  
 CC InterPro; IPR000569; HECT\_domain.  
 CC InterPro; IPR002349; WW.  
 CC Pfam; PF00168; C2; 1.  
 CC Pfam; PF00632; HECT; 1.  
 CC Pfam; PF00397; WW; 4.  
 CC PRINTS; PR00403; WWDOMAIN.  
 CC SMART; SM00233; C2; 1.  
 CC SMART; SM00119; HECTC; 1.  
 CC SMART; SM00456; WW; 4.  
 CC PROSITE; PS00499; C2\_DOMAIN\_1; FALSE\_NEG.  
 CC PROSITE; PS00004; C2\_DOMAIN\_2; FALSE\_NEG.  
 CC PROSITE; PS0237; HECT; 1.  
 CC PROSITE; PS01159; WW\_DOMAIN\_1; 4.  
 CC PROSITE; PS00020; WW\_DOMAIN\_2; 4.  
 KW Ub1 conjugation pathway; Ligase; Repeat.  
 FT DOMAIN 20 100  
 FT C2 DOMAIN.  
 FT DOMAIN 300 333  
 FT WW 1.  
 FT DOMAIN 330 363  
 FT WW 2.  
 FT DOMAIN 405 437  
 FT WW 3.  
 FT DOMAIN 444 477  
 FT WW 4.  
 FT HECT.  
 FT DOMAIN 536 870  
 FT BINDING 838 858  
 FT E -> K (IN REF. 1).  
 FT CONFLICT 136 136

FT	CONFLICT	394	395	SS -> FW (IN REF. 1).	
SEQ	SEQUENCE	870 AA;	98911 MW;	FCCD75CBA61F2204 CRC64;	
Query Match		31.5%;	Score 1225;	DB 1; Length 870;	
Best Local Similarity		35.5%;	Pred. No. 1.2e-81;		
Matches	281;	Conservative	115;	Mismatches 227; Indels 169; Gaps 19;	
QY	28	PFKIVDVGSGQCHSDTVKNTIDPKWQHY----	DLVYGKTSITISVNNHKKIHKQG	83	
Db	152	PNGSALTGGS-QLPSSDSSGTVAPE-NRHQPFSTNCFGGRS----	-----RTHRHS	198	
QY	84	AGFLGCVRLLSNAISRUKDGYQRDLCKLNPSDSDAVRGQIVSVLSQTRDR--	LGTTGGSV	141	
Db	199	AS-----	-----AETPATGQSGFARSRRQPVKNSGHS	228	
QY	142	VDGCGLLNKGTVYED-----	SGGRPLSCFMEE-----	PAPYTDSTGAAAGGN	186
Db	229	GLANGTVNDPEFTATDPEEPSVVGVTSPPAAPLSVTPNPNTTSLPAPATPAEG-----	-----	281	
QY	187	CRFVESQD--ORLQARLRNPDVRGSLQTPQNRPHGHQSPELPE-----	-----	230	
Db	282	---EPESTSGTQLP-----	---AAQAPDALPAGWEQRELPNGRVVYVDHNTKTTT	327	
QY	231	---GYEORTTVQGQVYFLHTQTGVSTWHDPRIFRDLN-----	-----	264	
Db	328	WERPLPGWEKRTDPRGRFYVVDHNRTRTTTWQRTAEYVRNVEQWSQRNLQAGMQHFS	387		
QY	265	---SVNCDLGLPLPGWEVRSYSGRIYFVDHNNRTQTDDP-----	-----	304	
Db	388	QRFLYQSSASTDHDPLGLPLPGWEKRDQ--NGRVVYVNHNRTRTQWEDPRPTQGMIOEPAL	446		
QY	305	---LHHMNHOCQLKEFSQPLPLPSEGSLEDEELPAQYERDLVQKLKVLRL	352		
Db	447	PPQWEMKYTSEGVRVYFVDHNRTRTTFKQPRPGESGTQKQSGPAYDRESFWKXHQFRFLC	506		
QY	353	HELSQLQPGACHCRIEVSREEIFEEYSRQITMKWRPKDLKRLMKVPRGEGLDYGVGARE	412		
Db	507	HSNAL-----PSHVKISVSROTLLFEDSFQQIMNMKPYDLRRRLYII	IMRGEGLDYGGIARE	562	
QY	413	WLXLLCHEMLNPYYGLFOYSTDNIYMLQINPDSSINPDHLUSYFHFVGRIMGLAVEHGYI	472		
Db	563	WFFLLSHEVLNPMYCLFEVAGKNYCLQINPASSINPDHLTYFRFTRGFAMALYHGKTI	622		
QY	473	NGGFTVPFYKQLGKPIQISDLSESVDPELHKSUVILENDITPV--LDHTFCVEHNAFGR	531		
Db	623	DTGFTLFPFKRMLNKRPPTLKLESIDPEFYNSIVTWIKENNLBECGLLELYPIQDMEILGV	682		
QY	532	LQHELKPNGRNVPTBENKEVYRLVYNNRMFEGIEAQFLAQKGFNELIPQHLLXFPDQ	591		
Db	683	TTHELKGGESIRVTENKEEYIMLHTDWRFTRGVBEQTKAFUDGNEVAPLEWLRYFDE	742		
QY	592	KELELIIGGLDKIDLNDKWSNTRLKECHVADSNTRVRFQAVETFDERRARLILQFVTGST	651		
Db	743	KELELMCGMQEIDMSDWQKSTIYRHVTKNSKQIQWFQVYVWEMDNKRIILILQFVTGTC	802		
QY	652	RVPLOQFKALQSGTAGRPLFTIHLIDANTDNLKPAHCTCFNRIDIPPVESYEKLYEKL	711		
Db	803	RUPVGGFAELIGSN---GPKFCIDIKVGKET--WLPSHSTCFNRLDLPYKSYEQLEKLL	858		
QY	712	TAVEETCGFAVE	723		
Db	859	YALEETEGFQGE	870		
RESULT 15					
NED4 MOUSE					
ID	NED4 MOUSE	STANDARD;	PRT;	887 AA.	
AC	P46935; O08758; Q8GB3;				
DT	01-NOV-1995 (Rel. 32, Created)				
DT	10-OCT-2003 (Rel. 42, Last sequence update)				
DT	10-OCT-2003 (Rel. 42, Last annotation update)				
GN	Ubiquitin-protein ligase Nedd-4 (EC 6.3.2.-)				
GN	NEDD4 OR NEDD-4 OR NEDD4A OR KIAA0093.				

CC	-!	FUNCTION: Involved in the embryonic development and differentiation of the central nervous system.
CC	-!	PATHWAY: Ubiquitin conjugation; third step.
CC	-!	SUBUNIT: Binds SCNN1A, SCNN1B and SCNN1G (By similarity).
CC	-!	Interacts with UBE2D2.
CC	-!	SUBCELLULAR LOCATION: Cytoplasmic.
CC	-!	TISSUE SPECIFICITY: Brain.
CC	-!	MISCELLANEOUS: A cysteine residue is required for ubiquitin-thiolester formation.
CC	-!	SIMILARITY: Contains 1 C2 domain.
CC	-!	SIMILARITY: Contains 3 WW domains.
CC	-!	SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL oustation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

```

ENBL; D85414; BAA12803.1; -.
DR ENBL; U96635; AAB63360.1; ALT_FRAME.
DR ENBL; AK088620; BAC40458.1; -.
DR ENBL; AK088767; BAC40558.1; -.
DR ENBL; AK122203; BAC65485.1; -.
DR HSSP; Q13526; IPIN.
DR MGD; MGI:97297; Nedd4.
DR GO; GO:0005829; Cytosol; IDA.
DR GO; GO:0000151; C:ubiquitin ligase complex; IPI.
DR GO; GO:0005155; F:protein binding; IPI.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
DR InterPro; IPR000008; C2.
DR InterPro; IPR008973; C2_CaLB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR001202; WW_Rfp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00337; WW; 3.
DR SMART; SM00239; C2.1.
DR SMART; SM00119; HECTC; 1.
DR SMART; SM00456; WW; 3.
DR PROSITE; PS00499; C2_DOMAIN_1; 1.
DR PROSITE; PS00004; C2_DOMAIN_2; 1.
DR PROSITE; PS00237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 3.
DR PROSITE; PS00020; WW_DOMAIN_2; 3.
DR Ub1 conjugation pathway; Ligase; Repeat.
FT DOMAIN 65 166 C2_DOMAIN.
FT DOMAIN 249 282 WW 1.
FT DOMAIN 405 438 WW 2.
FT DOMAIN 460 493 WW 3.
FT DOMAIN 552 587 WW 3.
FT BINDING 854 854 UBIQUITIN (BY SIMILARITY).
FT SEQUENCE 887 AA; 102705 MW; AE7DD3ED63986C50 CRC64;

```

	Query Match	31.5%;	Score 1225;	DB 1;	Length 887;
	Best Local Similarity	33.6%;	Pred. No. 1.3e-81;		
	Matches 291; Conservative	123;	Mismatches 238;	Indels 214;	Gaps 23;
Qy	7 IRLTVLCAKMLAKDPFLPPPAKIV----	VDSGGCHSDTDVKNLTLPKWGHY-----	58		
	: : :	:	:	:	:
	: : :	:	:	:	:
Dd	80 VRVKVIAGIGLAKKDILGASPDYPVRVTLYDFMSGILTYSVQTIKKSLNPKWNEILLFRV		139		
Qy	59 -----	DLYGKT-D-SITISVMN-----	HKHHKQG 83		
	:	:	:	:	:
Dd	140 LPQRHRLFVEFDENLRTRDDFLGQVDPLPPLPTENPRMERPYTFKDVLHPRSHKSrv		199		
vq	84 AGFLGCVRLLLSNAISRDKDT-----		103		

200	QY	----	RLKMTYLPKNGSEADQAELEBPGWVVLDPQDAATHLHPPEP	244
104	QY	----	GY-QRLDLC-----KLPSTDAVRGQIVSVLQTRDI	135
249	Db	SPLP	PGWEERQDVLGRTRYVYNHESRKTQWKRPSPDDLTDEDDMDQLQAFATTRRQI	308
136	QY	GTGG	VVDCRGLLEN-----EGTVYEDSGRPLSCFME-----EP	171
309	Db	SEDV	GDPNRESPENWEIVREDENTEVSGOAVQSPSGCHIDVQTHLAEENTRLAVCGNP	368
172	QY	A----	PYTDSTGAAAGG--NCRFVESPQORLOAQRLNPDVRGSLQTPQNRPHGHQSP	226
369	Db	ATSP	QVITSSNHSSRGGSLQTCIFEEQPTLPVLLPT-----SS	405
227	QY	ELPG	YEQRTTVQGVQVFLHTQTGVSTWHPD-----RIPRDL-NSVNCDELPLPPG	277
406	Db	GUPP	GEWEKQDGRSYYVDHNSKTTTWSKPTQDDPRSKI PAHLGKTDSDNDLGLPLPG	465
278	QY	MEVRS	TVSGRIYFYVDHNNRTTQTDPRLHHIMNHQCOLKEPSOPLPLPSEGSLEDEELPA	337
466	Db	NEERT	HTDGRVFFINHNKIKTQWEDPQLQV-----AITGPAVP-----	504
338	QY	QRYER	DLVOKLVLRHLSLQOQAGHCRTEVSEERIFEEBSYQIMKMRPKD-LKXRLMV	396
505	Db	YSDY	KRKYEFFRKLKKQTDIPNKFWKLRANILEDSYRRIMGVXRADLLKARLWI	562
397	QY	KFRGE	GLDYGAVAREWLYLLICHEMLNPPYGLFYQSTDNIYMLQINPDSSI-NPDHLSYF	455
563	Db	EDGE	KLDYGAVAREWFLISKEMFNPPYGLFEYSATDNYTTQINPNSGLCNDEHLSYF	622
456	QY	HFVG	IRMLAVPHGYHNGFTVPFYKQLLGKPIQLSDLESVDPELHKSVLWILENDIIP	515
623	Db	KFIG	VAGVYHVKGLLDGFFIRFPYQOMLOKUITLHDMSVDSYSSURLEWLENDPTE	682
516	QY	VLDT	HFCEVHNAFRILLOHELKPNRNVVPTENKXEYVRLYVNNRPFMGIEAQFLALOK	575
683	Db	LDLAF	IDLELFGQTHOHELKTCGSLIVVTNNKXEYIVLVQWRFVNRIOKMAAFKE	741
576	QY	GFNEL	IPOHLKPFDOKELELIIGGLPKIDLNWKSNTRLKHCUA-DSNIVRFWQAVET	634
742	Db	GFPELL	IQDLIKIFENELELIMCGLGDVDVNDWREHTKYKNGYSMNHOVHWFKAVNM	801
635	QY	PDERR	ARLLQFVTSRVPLQGFKALQGSTGAAGPRLFTIHLIDANTONLPKAHTCFNR	694
802	Db	NDSER	IRLLQFVTGTSRVPMNGFAELYGSN---GPQSFVEQW-GTPDKLPRAHTCFNR	857
695	QY	IDIPP	YESYKLYEKLTLTAVEETCGF	720
858	Db	LPLPY	ESFDELWDKLOWAIENTOGF	883

Search completed: September 21, 2004, 07:44:16  
Job time : 24.7111 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:36:41 ; Search time 87.9789 Seconds  
(without alignments)  
2592.890 Million cell updates/sec

Title: US-10-009-945-2  
Perfect score: 3884  
Sequence: 1 GGSSKIRLTVCAXNLAKK.....EKLYEKLTAVEETCGFAVE 723

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPREMBL 25:\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3792.5	97.6	728	11	Q8K300 mus musculu
2	2206	56.8	1061	5	Q9W853 drosophila
3	1895	48.8	355	11	Q8BSC0 mus musculu
4	1559.5	40.2	295	4	Q8NDDB8 homo sapien
5	1341	34.5	786	3	O14326 schizosacch
6	1340.5	34.5	288	4	Q96DE7 homo sapien
7	1269	32.7	911	4	Q8N5A7 homo sapien
8	1247	32.1	955	4	Q95PU5 homo sapien
9	1245	32.1	947	4	Q7Z5F1 homo sapien
10	1245	32.1	975	4	Q7Z5N3 homo sapien
11	1243	32.0	967	4	Q7Z5F2 homo sapien
12	1241.5	32.0	820	4	Q9NT88 homo sapien
13	1239	31.9	949	5	Q9Y0H4 drosophila
14	1228.5	31.6	855	4	Q8WU09 homo sapien
15	1228.5	31.6	858	4	Q8W58 homo sapien
16	1228.5	31.6	971	13	O42573 xenopus lae

17	1228.5	31.6	995	4	O43165	O43165 homo sapien
18	1226	31.6	854	4	Q9H2W4	Q9H2W4 homo sapien
19	1224	31.5	855	11	Q8CF10	Q8CF10 mus musculu
20	1223.5	31.5	835	11	Q8BRT9	Q8BRT9 mus musculu
21	1215	31.3	855	11	Q99PK2	Q99PK2 mus musculu
22	1214.5	31.3	834	5	Q95R64	Q95R64 drosophila
23	1213	31.2	794	5	Q9N2Z7	Q9N2Z7 caenorhabdi
24	1210	31.2	792	5	O95XU3	Q95XU3 caenorhabdi
25	1207.5	31.1	376	13	Q98RS4	Q98RS4 xenopus lae
26	1200.5	30.9	258	11	Q9CSE3	Q9CSE3 mus musculu
27	1196.5	30.8	956	5	Q8IQR6	Q8IQR6 drosophila
28	1195.5	30.8	518	5	Q8TOC8	Q8TOC8 drosophila
29	1191	30.7	1007	5	Q9VVI3	Q9VVI3 drosophila
30	1186	30.5	222	11	Q80TB1	Q80TB1 mus musculu
31	1151	29.6	838	5	Q95TQ0	Q95TQ0 drosophila
32	1129	29.1	724	5	Q9BKW4	Q9BKW4 caenorhabdi
33	1026.5	26.4	671	3	Q9UTG2	Q9UTG2 schizosacch
34	1002	25.8	1585	4	Q9HCC7	Q9HCC7 homo sapien
35	989	25.5	1581	4	Q9P2P5	Q9P2P5 homo sapien
36	968.5	24.9	1583	11	Q8K4P8	Q8K4P8 mus musculu
37	896.5	23.1	437	5	O9W3Z5	Q9W3Z5 drosophila
38	886	22.8	339	11	Q8BIA6	Q8BIA6 mus musculu
39	859	22.1	1906	4	O150Z9	O150Z9 homo sapien
40	859	22.1	3360	4	Q8NG67	Q8NG67 homo sapien
41	854.5	22.0	403	11	Q8VD14	Q8VD14 mus musculu
42	854.5	22.0	444	11	Q7TMY8	Q7TMY8 mus musculu
43	854.5	22.0	1080	11	Q921M5	Q921M5 mus musculu
44	854.5	22.0	1484	11	Q8CPH2	Q8CPH2 mus musculu
45	854.5	22.0	4065	3	Q9P4Z1	Q9P4Z1 neurospora

ALIGNMENTS

RESULT 1  
Q8K300 PRELIMINARY; PRT; 728 AA.  
ID Q8K300  
AC Q8K300;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DE Similar to E3 ubiquitin ligase SMURF1.  
GN 4930431E10RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
DR EMBL; BC029097; AAH29097.1; -;  
DR MGD; MGI:1923038; 4930431E10RIK.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2, CaLB.  
DR InterPro; IPR000569; HECT domain.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 2.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECTC; 1.  
DR SMART; SM00456; WW; 2.  
DR PROSITE; PS00499; C2 DOMAIN 1; 1.  
DR PROSITE; PS50004; C2 DOMAIN\_2; 1.  
DR PROSITE; PS50237; HECT; 1.  
DR PROSITE; PS01159; WW DOMAIN 1; 1.  
DR PROSITE; PS50020; WW DOMAIN\_2; 2.

KW	Sequence	728 AA	83083 MW	C355291B9D8AD757 CRC64
QY	2 GSSIKIRLTVCANLAKKDFRLPDPRAKIVVDSGGCHSTDTVNTKLDPKNCHVDLY	61		
Db	10 GSSIKIRLTVCANLAKKDFRLPDPRAKIVVDSGGCHSTDTVNTKLDPKNCHVDLY	69		
QY	62 VGTDSITISVNNKKIKKQAGAGLGCVRLLSNAISLKDGTQYRLDCLCKLNPSDSDAV	121		
Db	70 VGTDSITISVNNKKIKKQAGAGLGCVRLLSNAISLKDGTQYRLDCLCKLNPSDSDAV	129		
QY	122 RGOIVVSLQTRDRIGTGSVWDCRGLLENEGTVVDSGPRPLSCFWEPEAPYTDSTGAA	181		
Db	130 RGOIVVSLQTRDRIGTGSVWDCRGLLENEGTVVDSGPRPLSCFWEPEAPYTDSTGAA	189		
QY	182 AGGNCRCFVSPSQDRLQAOALRNPVGRSLQTPQNRPHGHPSPPEGYEQRTTVOGQ	241		
Db	190 AGGNCRCFVSPSQDRLQAOALRNPVGRSLQTPQNRPHGHPSPPEGYEQRTTVOGQ	249		
QY	242 VFLHTQTVGTSVWHPDRIPRDLNSVNCDELGPLPGWEVRSTVSGRIYFVDHNNRTTQFT	301		
Db	250 VFLHTQTVGTSVWHPDRIPRDLNSVNCDELGPLPGWEVRSTVSGRIYFVDHNNRTTQFT	309		
QY	302 DPLHMHMHQCOLKEPSQPLPSEGSLEDELPAPRYERDLVOKLVLRLHLSLQOPQ	361		
Db	310 DPLHMHMHQCOLKEPSQPLPSEGSLEDELPAPRYERDLVOKLVLRLHLSLQOPQ	369		
QY	362 AGHCRLEVSREIFESYRQIMKMPKDLKRLMKVKGREGDLYGGVAREWLYLLCHEM	421		
Db	370 AGHCRLEVSREIFESYRQIMKMPKDLKRLMKVKGREGDLYGGVAREWLYLLCHEM	429		
QY	422 LNPYGLFOYSDNTIYMLQINPDSSINPDHLSYHFVGRIMGLAVFRGHVINGGFTVPFY	481		
Db	430 LNPYGLFOYSDNTIYMLQINPDSSINPDHLSYHFVGRIMGLAVFRGHVINGGFTVPFY	489		
QY	482 KOLLGKPIQLSDLESVDDELKSLVWILENDITPVLDTHTFCVFNAGRIQLHLEKNGR	541		
Db	490 KOLLGKPIQLSDLESVDDELKSLVWILENDITPVLDTHTFCVFNAGRIQLHLEKNGR	549		
QY	542 NVPVTEENKKEYRVLVYNNRMFGIEAQFLAQKFNELIPQHLKLPFDQKELELIIGL	601		
Db	550 NVPVTEENKKEYRVLVYNNRMFGIEAQFLAQKFNELIPQHLKLPFDQKELELIIGL	609		
QY	602 DKIDLNDKSNRLKHCVDNSIVWFQWAVETDEERRARLLQFVTGSTRVPLQGFAL	661		
Db	610 DKIDLNDKSNRLKHCVDNSIVWFQWAVETDEERRARLLQFVTGSTRVPLQGFAL	669		
QY	662 QGSTGAAGRLFTIHLIDANTDNLPAKTCFNRIDIPYSEYKLYEKLTAVEETCGFA	721		
Db	670 Q---GAAGRLFTIHLIDANTDNLPAKTCFNRIDIPYSEYKLYEKLTAVEETCGFA	726		
QY	722 VE 723			
Db	727 VE 728			

## RESULT 2

ID	Q9V853	PRELIMINARY	PRT	1061 AA
AC	Q9V853	Q9V853		
DT	01-MAY-2000	(TRENDELrel. 13, Created)		
DT	01-OCT-2003	(TRENDELrel. 22, Last sequence update)		
DE	CG4943 protein (E3 ubiquitin ligase) (Ubiquitin-protein ligase) (Smad-			
DE	ubiquitin E3 ligase (Smurfin))			
GN	LACK OR SMURF1 OR CG4943			
OS	Drosophila melanogaster (Fruit fly)			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			

Ephydroidea; Drosophilidae; Drosophila.

OC [1]  
CX NCBI\_TaxID=7227;  
RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherz S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cline S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Benson J., An H., Baldwin D., Benzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuncio J.,  
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of Drosophila melanogaster genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Mira S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminiker J.S., Prochnik S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of Drosophila melanogaster genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]

RP SEQUENCE FROM N.A.  
RA Adams M.D., Ceiniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Laurencon A., Hawley S.;  
RL "Molecular cloning of a type E3 Ubiquitin ligase";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RA Pedos S.D., Hanson K.K., Wang Y.-C., Ferguson E.L.;  
RL "The Dsmurf ubiquitin-protein ligase restricts BMP signaling spatially  
RT and temporally during Drosophila development";  
RL Dev. Cell 1:0-0(2001).  
RN [8]  
RP SEQUENCE FROM N.A.  
RA Liang Y.-Y., Lin X., Feng X.-H.;  
RL "dSmurf1, a Smad-ubiquitin E3 ligase, specifically targets dpp-  
RT activated Wad protein for degradation";  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
DR EMBL; AR003802; AAF57824.3; -;  
DR EMBL; AF216521; AAF21125.1; -;  
DR EMBL; AF416571; AAL09691.1; -;  
DR EMBL; AF464851; RAM09646.1; -;  
DR HSSP; Q13526; 1PIN.  
DR FlyBase; FBgn0029006; lack.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0016874; F:ligase activity; IEA.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
DR InterPro; IPR000008; C2\_DOMAIN.  
DR InterPro; IPR008973; C2\_calB.  
DR InterPro; IPR000569; HECT\_domain.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 3.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00456; WW; 3.  
DR PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
DR PROSITE; PS02337; HECT; 1.  
DR PROSITE; PS01159; WW\_DOMAIN\_1; 2.  
DR PROSITE; PS00020; WW\_DOMAIN\_2; 3.  
KW Ligase.  
SQ SEQUENCE 1061 AA; 115675 MW; 6BBCC550F5129163 CRC64;

Query Match  
Best Local Similarity 43.0%; Score 2206; DB 5; Length 1061;  
Matches 454; Conservative 108; Mismatches 148; Indels 346; Gaps 14;

QY 6 KRLIVLCAKNAKDFRLPPFAKIVVDGSGQCHSTDTVKNLDPKNOHYDLYGKT 65  
DB 14 KVRITILCAKNAKDFRLPPFAKIVVDGSGQCHSTDTVKNLDPKNOHYDLYGKT 73  
QY 66 DSITTSVNNKHKKQAGFLGCVRLLSNAISRLKDTGYQFLDLCKLNPSDPAVRQI 125  
DB 74 DAITITVNNQRIKH--GSGFLGCVRIPAFNISQIKGAGFQLDLGKLSPPDDELVRQI 131  
QY 126 VVSLOTDRIRIGT-----GSVVDRCGLENE-----GTIVY----- 155  
DB 132 IISLLSKDGPSSGNPLAIVGSGDVRGPDSEDDSDSLPEGWEERTDNGRVYVNHATK 191  
QY 156 -----EDSGFGRFLSCFMEEPAPYTDSTGAAGGNCRFVE----- 191

Db 252 SVTASDERHSHSTELSSVGKENTSPPTPVSATTTTPGKKTSSNSSSAGG--RTLEORPT 308  
QY 192 ----SPSQDQRLQACRLNPDVRGSLQTPQNRPHGHOSPE----- 227  
Db 309 NEPATPSSITTSASVRLHSND--NHVKTPRHQTNGHAPPESTPTSPGQQQYVNGNAQNG 366  
QY 228 ----- 227  
Db 367 STSGNGSQAAQPOSASNGWTQEDAAITTSPTTSPRHSQSPPPTNISPPASVTPSAN 426  
QY 228 ----- 227  
Db 427 GNHSPNANSTPAGSGGSGRSYTAATPCQSRQSRSSRQCGEESSTRSSRGTNGTSG 486  
QY 228 -----LPEGYEQRITVQGVYFLHTQTGVSTWHPRI 259  
Db 487 GGGGGGQRYASAAIAAQAARPFDLDPGYEMRTTQQGVYFYHIPTGVSTWHPRI 546  
QY 260 PRDLNS--VNCDELGLPPGKWEVSTVSGRIYFVDHNNRTQFTDPRLHHM----- 309  
Db 547 PRDFTQHLTLDAIGPLPSGWEQKTAGRVYFVDHNNRTQFTDPRLSGILQMIIRGT 606  
QY 310 ---NHQCQLKEPSQFLPLPS-----EG 328  
Db 607 VPPTSAAAGTAPPSATPATAAATAVPQATPASNATTTTLTTTNPPIRIVPDLPGQ 666  
QY 329 SLEDEELPAQRYERDLVOKLVLEHLSLQOQAGHCRIEVSRBEIEFESYRQIMQWRPK 388  
Db 667 LLEGADL-LPKYRRDLVGLKRALRTELTQMQPOSGHCRLEVSRENEIFESYRLIMKRAK 725  
QY 389 DLKRLVMKPRGEBGLDYGVAEWLYLLCHEMLNPVYGLFOYSTDNVLMQINPDSSIN 448  
Db 726 DMKRLVMKPRGEBGLDYGVAEWLYLLCHEMLNPVYGLFOYSTDNVLMQINPDSSIN 785  
QY 449 PDHLSYFHFVGRINGLAVFHHYINGGFTVPFYKQLGKPIQLSDLESVDPFLHKSILWI 508  
Db 786 PDHLSYFHFVGRITGLIAVFGHCLDGGFTTFFYKQLLNKLTILGDIQGVDPDLHRSLTWM 845  
QY 509 LENDITPVLDTFCVHNAGRILOHELKPNRNPVTEENKKEYVLYYNNWFMRIEZA 568  
Db 846 LESNISGIIESTFVENNSFGALVVHELKPGGASIPVTEENKKEYVLYYNNWFMRIEQA 905  
QY 569 QFLALQKGFNELIPQHLLKPFQKLELIIGGLDKIDLNDWKSNTRLKHCVADSNIVRWF 628  
Db 906 QFLALQKGFCELIPSHLLRPFDERELVIGGSISSIDVNDWRNTRLKHCTNETTQVLMWF 965  
QY 629 WQAVETDEERRARLLQFVTGSTRVPLQGPALQGSTGAAGPRLFTIHL-IDANTDNLPK 687  
Db 966 WQVYESYSSEMRARLLQFVTGSSRVPLQGFALQGSTGAVGPRFLTHTADVPTQNLPK 1025  
QY 688 AHTCFNRIDIPPPYESYEKLYELLTAVEETCGFAVE 723  
Db 1026 AHTCFNRIDIPPPYETVQLLCKLTQAVEETCGFAVE 1061

RESULT 3  
Q8BSC0  
ID Q8BSC0 PRELIMINARY; PRT; 355 AA.  
AC Q8BSC0;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical HECT domain (Fragment).  
GN 4930431E10RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_faxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 DR EMBL; AK034736; BAC28813.1; -;  
 DR MGI; 1923038; 4930431B10Rik.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0004942; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000569; HECT\_domain.  
 DR Pfam; PF00632; HECT; 1.  
 DR SMART; SM00119; HECTC; 1.  
 DR PROSITE; PS50237; HECT; 1.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 355 AA; 41315 MW; 93B39B4C82F86DD1 CRC64;  
 Query Match 48.8%; Score 1895; DB 11; Length 355;  
 Best Local Similarity 99.7%; Pred.No. 3.7e-147;  
 Matches 354; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 369 VSREIFEESVQIMKMRPKDKRLMKVFRGEGGLDYGVAREWLYLLCHEMLNPFYGL 428  
 Db 1 VSREIFEESVQIMKMRPKDKRLMKVFRGEGGLDYGVAREWLYLLCHEMLNPFYGL 60  
 QY 429 FQYSTDNIMQLNPDSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQLLGP 488  
 Db 61 FQYSTDNIMQLNPDSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQLLGP 120  
 QY 489 IQLSDLESVDPELHKSLVWILENDITPVLDHTFCVEHNAFGRILQHLKPKDQKELELIIGGLDKIDL 548  
 Db 121 IQLSDLESVDPELHKSLVWILENDITPVLDHTFCVEHNAFGRILQHLKPKDQKELELIIGGLDKIDL 180  
 QY 549 NKKEYVRLVYVNRVFRMGIEAQFLALQKGFNEIIPQHLLKPFDPQKELELIIGGLDKIDLND 608  
 Db 181 NKKEYVRLVYVNRVFRMGIEAQFLALQKGFNEIIPQHLLKPFDPQKELELIIGGLDKIDLND 240  
 QY 609 WKSNTLRKHCVDADSNVRFVQAVETFEERRARLLQFVTGSTRVPLQGFALQSGTGAA 668  
 Db 241 WKSNTLRKHCVDADSNVRFVQAVETFEERRARLLQFVTGSTRVPLQGFALQSGTGAA 300  
 QY 669 GPRLFTIHLIDANTDNLPAKHTCFNRIDIPPYSEYEKLYEKLITAVEETCGFAVE 723  
 Db 301 GPRLFTIHLIDANTDNLPAKHTCFNRIDIPPYSEYEKLYEKLITAVEETCGFAVE 355  
 RESULT 4  
 Q8NDD8 PRELIMINARY; PRT; 295 AA.  
 ID Q8NDD8  
 AC Q8NDD8  
 DT 01-OCT-2002 (TREMELrel. 22, Created)  
 DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN DKFZP564H223.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Tissue=Brain;  
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL834242; CAD38919.1; -;  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000569; HECT\_domain.  
 DR Pfam; PF00632; HECT; 1.  
 DR SMART; SM00119; HECTC; 1.  
 DR PROSITE; PS50237; HECT; 1.

KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 295 AA; 34030 MW; 0507325127A943EA CRC64;  
 Query Match 40.2%; Score 1559.5; DB 4; Length 295;  
 Best Local Similarity 98.7%; Pred.No. 1.1e-119;  
 Matches 293; Conservative 1; Mismatches 0; Indels 3; Gaps 1;  
 QY 427 GLFOYSTDNIMQLNPDSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQLLGP 486  
 Db 2 GLFOYSTDNIMQLNPDSINPDHLSYFHFVGRIMGLAVFHGHYINGGFTVPFYKQLLGP 61  
 QY 487 KPIQLSDLESVDPELHKSLVWILENDITPVLDHTFCVEHNAFGRILQHLKPKDQKELELIIGGLDKIDL 546  
 Db 62 KPIQLSDLESVDPELHKSLVWILENDITPVLDHTFCVEHNAFGRILQHLKPKDQKELELIIGGLDKIDL 121  
 QY 547 BENKKEYVRLVYVNRVFRMGIEAQFLALQKGFNEIIPQHLLKPKDQKELELIIGGLDKIDL 606  
 Db 122 BENKKEYVRLVYVNRVFRMGIEAQFLALQKGFNEIIPQHLLKPKDQKELELIIGGLDKIDL 181  
 QY 607 NDWKSNTLRKHCVDADSNVRFVQAVETFEERRARLLQFVTGSTRVPLQGFALQSGTG 666  
 Db 182 NDWKSNTLRKHCVDADSNVRFVQAVETFEERRARLLQFVTGSTRVPLQGFALQSGTG 238  
 QY 667 AAGPRLFTIHLIDANTDNLPAKHTCFNRIDIPPYSEYEKLYEKLITAVEETCGFAVE 723  
 Db 239 AAGPRLFTIHLIDANTDNLPAKHTCFNRIDIPPYSEYEKLYEKLITAVEETCGFAVE 295  
 RESULT 5  
 O14326 PRELIMINARY; PRT; 786 AA.  
 ID O14326  
 AC O14326  
 DT 01-JAN-1999 (TREMELrel. 09, Created)  
 DT 01-JAN-1999 (TREMELrel. 09, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Ubiquitin--protein ligase.  
 GN SPBC1689.11C.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972h-;  
 RA Volckaert G., Wood V., Rajandream M.A., Barrell B.G.;  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 CC 1- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
 DR EMBL; Z99759; CAB16903.1; -;  
 DR FIR; T39585; T39585.  
 DR HSSP; Q13526; IPIN.  
 DR GeneDB\_SPombe; SPBC1689.11C; -;  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0016874; F:ligase activity; IEA.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.  
 DR GO; GO:0006512; P:ubiquitin cycle; IEA.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR008973; C2\_CalB.  
 DR InterPro; IPR000569; HECT\_domain.  
 DR InterPro; IPR002349; WW.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00632; HECT; 1.  
 DR Pfam; PF00397; WW; 3.  
 DR PRINTS; PR00403; WWDOMAIN.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00119; HECTC; 1.  
 DR SMART; SM00456; WW; 3.  
 DR PROSITE; PS00439; C2\_DOMAIN\_1; 1.  
 DR PROSITE; PS50034; C2\_DOMAIN\_2; 1.  
 DR PROSITE; PS50237; HECT; 1.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 3.

OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	[1]
RN	NCBI_TaxID=9606;
RP	SEQUENCE FROM N.A.
RC	TISSUE=Lung;
RA	Strausberg R.;
RL	Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; BC009527; AAH09527.1; -.
DR	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0016874; F:ligase activity; IEA.
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.
DR	InferPro; IPR000569; HECT_domain.
Pfam	Pf00632; HECT; 1.
SMART	SM00119; HECTC; 1.
PROSITE	PS50237; HECT; 1.
KW	Ligase.
FT	NON TER
SQ	SEQUENCE 288 AA; 33255 MW; PE2B43E300B66537 CRC64; Query Match 34.5%; Score 1340.5; DB 4; Length 288; Best Local Similarity 84.8%; Pred.No.1e-101; Matches 246; Conservative 23; Mismatches 18; Indels 3; Gaps 1;
Qy	434 DNIYMLQINPDSNPDLHSVEHFVGVGIMGLAVPHGHVINGGFVPFYKOLIGRPIQLSD 493           :           :           :           :           :           :           : Db 2 DDITLTQINPDSPANVPHLSFIFFHVGRIMGVMVFHGHIYDGGFTLPFYKOLLGKSITLDD 61
Qy	494 LESVDPELHKSLVMILENDITPVLDHTFCVEHNAGFRILQHLEKPNGNRPVTENKKKEY 553           :           :           :           :           :           :           : Db 62 MELVDPDLHNSLVMIENDITGVLDHTFCVEHNAYGEIIHQELKPCKSI PVNEENKEY 121
Qy	554 VRIYVWRFMRGTEAQFALQKGFNELIPHLIKPPDOKELELIIGGLDKIDLMDWKSNKT 613           :           :           :           :           :           :           : Db 122 VRUYVWRFLURG-EAQFALQKFNEVIPHLLKTFDEKELELIICGLGKIDVNDWKVNT 181
Qy	614 RLKHCVADSNIVFWFOAVTPTDEERRARLLQVGTSTRVLQGFKALOGSTGAAGRPLF 673           :           :           :           :           :           :           : Db 182 RLXHCTPDSNIVKWFWKNVFPFEERARLLQVTSRRVPLOGFKALO--GAAGRPLF 239
Qy	674 TIHLIDANTNLPKATCFNRIDIPPYESYEKIYEXLLTAVERTCGFAVE 723           :           :           :           :           :           :           : Db 239 TIHQIDACTNNLKPAHTCFNRIDIPPYESYEKIYEXLLTAIEETCGFAVE 288
RESULT 7	
Q8NSA7	PRELIMINARY; PRt; 911 AA.
ID AC Q8NSA7	
DT 01-OCT-2002	(T-EMBLrel. 22, Created)
DT 01-OCT-2002	(T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2003	(T-EMBLrel. 25, Last annotation update)
DE	Similar to neural cell expressed, developmentally down-regulated 4-like.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	[1]
RN	NCBI_TaxID=9606;
RP	SEQUENCE FROM N.A.
RC	TISSUE=Melanoma;
RA	Strausberg R.;
RL	Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
CC	-/- SIMILIARTY: CONTAINS 1 C2 DOMAIN.
DR	EMBL; BC032597; AAH32597.1; -.
DR	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.
DR	InferPro; IPR000008; C2.
DR	InferPro; IPR008973; C2_CalB.
DR	InferPro; IPR000569; HECT domain.



Query Match	32.1%	Score 1247;	DB 4;	Length 955;
Best Local Similarity	31.9%	Pred. No. 3.1e-93;		
Matches 310;	Conservative 125;	Mismatches 246;	Indels 292;	Gaps 24;
QY 2	GSSIKIRLTVLCAXNLAKKQFFRLDPFPAKI	---VVDGSGQCH---STDTKVNTLDLPKWQ	56	
DB	17 GESRILRVKVVSGIDLAKKQIFGASDPVVKLSLVADENRELALVQTKTIKTLINPKWNE	76		
QY 57	HYDLVYCKTD-SIIISVYNNHKKHKKGAGFLGCVRLLSNAI	---97		
DB	77 EFYFRVNPNSHRLLEFVFDENLRTRDD	---FLGQVDVPLSHLPDTEMTMERPTFKDFLL	133	
QY 98	-----SRLKDTGYQRLDCLKNLPSDIDAVRGQIVVSLQTRDRIGTCGVSVDRCG	---146		
DB	134 RPRSHKSRVK--GFLRLKMAVMPKNG	---GQDEENSQDQDDMEHCWVVDSDNSAQH	186	
QY 147	-----LLENGGVY	---155		
DB	187 QBELPPPLPPGWEKVDNLRGTYYVNNHNRRTQWHRPMLDMVSESDNNIRQINQEAH	246		
QY 156	-----EDSGP	---GRPLSCFMEEPAPTYDSTGAA	---AGGNCREFVES	192
DB	247 RFRFSRAHISEDLEPEPESEGDPPEPWETISEEVIAGDSLGLALPPPPASPGSRTPSQE	306		
QY 193	PSOD	---QRLQAORLRN	---206	
DB	307 LSELSSRLQITPDSNCEQFSSLIQRPSPRLRSCSVTDVAEGQLHPPSPSVAYVHTTPG	366		
QY 207	-----	---PDVR	---210	
DB	367 LPSGWEERKDAKGRYYVNNHNRRTTWTTRPIQLAEDGASGASATNSNNHLIEIQIRAPRS	426		
QY 211	-----	---GSLQTPQNRPHGHQSP	---LPEGYEQR	235
DB	427 LSSPTVTLAPLSEAKDSPVRAVKQDTLNPQSPQSPSPYNSPKQHKVQTSQTFPPGWEHR	486		
QY 236	TTVQGVYFHTQTGVSTWHDPRIP	---RDLSNVNCDLGLPPLPGWEYRSTVSGRIYF	290	
DB	487 IAPNGRPFPTDHTKTITWTEDPLKFPVHWKSTSLNPDLGLPLPGWEERIHLDCRTFY	546		
QY 291	VDHNNRTQTFDRLHHIMNHQCLKEPSQPLPLPSEGSLEDEELPAQRYERDLVQKLKV	350		
DB	547 IDNSKITQWEDPLQV	---PAITG	---FAPVYSREFKOKYDY	583
QY 351	LRHELSQLQPAQCHCRIEVSREBIFESYRQIKVM-RPKDLKKRLMVKFRGEBGLDYG	409		
DB	584 FRKLLAKPADIPNRFEMLKLRHNNIFESYRIRMSVKRPDLKARLWTEPSEKGLDYG	643		
QY 410	AREWLVLICHEMLNPPYGLFCYSTDTNMLQINPDSSI-NPDHLSYFHFVGRIMGLAVPH	468		
DB	644 ARWFFLLKEMFNPPYGLFEYSATNYTLQIINPNSGLCNEDHLSFTFTFGRVAGLAVPH	703		
QY 469	GHVINGFTVPFYKQLGKPIQLSDLESVDPELHLSLWILENDITPVLDPHTFCVEHNAF	528		
DB	704 GKLLDGFTRFPFYKMLMGKQITLNDMESVDSEYNSLKWILENDPTE-LDLMFCDIENF	762		
QY 529	GRILQHELKNGRMVPTENKKEYVRLVYNWFMERGIEAOFALOKGNEFLTPOHLLKP	588		
DB	763 GQTYQVDLPNGSEIWTNENKREYIDLVIQWFRVNRVQKMAFLEGFTTELLPILIKI	822		
QY 589	FDQKLELIGGDKIDLNDKWSNTRLKHCVADSN-IVRFWQAVETFDSEERARLLQFV	647		
DB	823 FDSNELELLMCGLDGVDVNDWRGHSYKNGCYCPNHPIQWFWKAVLLMDAEKRIALLQFV	882		
QY 648	TGSTRVPLQGFKALQSGTGAAGPRLTILHLIDANTDLNPKAHTCFNRIDIPPIPVSEKLY	707		
DB	883 TGTSRVPNGFAELYGSN---GQPLTTIEQW-GSPKLPRAHTCFNRLLDPPYETEDLUR	938		
QY 708	EKLLTAVETCGF	720		
DB	939 EKLLMAVENAQGF	951		

RESULT 9  
 Q7Z5F1 PRELIMINARY; PRT; 947 AA.  
 ID Q7Z5F1 AC Q7Z5F1; DT 01-OCT-2003 (TEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)  
 DE Uridiquitin ligase NEDD4h.  
 GN NEDD4L.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Qi H., Labrie C.;  
 RT "New splicing isoform of human Nedda-2.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY112985; AAM76730.1; -  
 KW Ligase.  
 SQ SEQUENCE 947 AA; 109429 MW; 95CB0FDEDCB96639 CRC64;  
  
 Query Match 32.1%; Score 1245; DB 4; Length 947;  
 Best Local Similarity 31.9%; Pred. No. 4.5e-93;  
 Matches 310; Conservative 125; Mismatches 246; Indels 292; Gaps 24;  
  
 QY 2 GSSIKIRLTVLCAKNLAKDFFRLPDPFAKI---VVDGSGOCH--STDVTXNTLDPKWQ 56  
 DB 9 GESILRVKVVSGIDLAKXDI FGASDPYVKLSYVPDENREIALVOTKI KKITLNPAWNE 68  
  
 QY 57 HYDLYVGKTD-STISVNNHKKLHKQAGFTGCVLLSNAI----- 97  
 DB 69 EFYFRVNPNSHRLFFVFDENLRTRD---FLQVDVPLSHLPTEDPTMERPYTFKDFLL 125  
  
 QY 98 -----SRLKDTGYQLDLCKLNPSTDVAVRQIVVSLQTRIGTGGSVNDCRG----- 146  
 DB 126 RFASHKSRVK--GFLRLKWAYPKNG-----GQDEENSQORDMHRGNEVVDNSASQH 178  
  
 QY 147 -----LLENEGTVY----- 155  
 DB 179 QEELPPPLPPGWEEKVDNLGRITVYVHNHNRITQWHRPSLMDVSSSDNNIRQINQEAH 238  
  
 QY 156 -----EDSGP-----GRPLSCFMEEPAPYTDSTGAA-----AGGNCRFVES 192  
 DB 239 RFRSRRH;SEDLPEPESGQDVPPEWTISIEVNIAGSLGLALPPPPASPGSRTSPOE 298  
  
 QY 193 PSCD-----QSLQACLRN----- 206  
 DB 299 LSELRLRLQITPDSNGEQFSSLI QREPSRLRSCSVTDVAEQGLHPPPSVAYVHTTGG 358  
  
 QY 207 -----PDVR----- 210  
 DB 359 LPSGWEERKDAKRTVYVHNHNRITTTWTRPIMQLAEDGASGATNSNNHLIPQIRPRS 418  
  
 QY 211 -----GSLCTPNQRPHGHOSPE-----LPGEYOR 235  
 DB 419 LSSPTVTLSPALEGAKDSPVRPAVKDTLSNPQSPSPINSKPKQHKVTSQSLPPGWEHR 478  
  
 QY 236 TTVOGVYFLHTQTGVSTWHDPRIP-----RDLSVNCDELGPLPGWEVRSTVSGRIYF 290  
 DB 479 IAPNGRPFFIDNTKTTTWEDPRLKFPVHMRSKTSILNPNDLGPLPGWEERIHLDRTFY 538  
  
 QY 291 VDHNRNTQFTDPRLHHIMNHOCQLKEFSQPLPSPESGSEDEELPAQRYERDLVQKLKV 350  
 DB 539 IDHNSKITQWEDPRIQN-----FAITG-----PAVYSEFQKQYD 575  
  
 QY 351 LAHELSLQOPQAGHCRIEVSRBEI FEESYRQIMKX-RPKDCLKRLMKVFRGSEGLDYGV 409  
 DB 576 FRKLLKPADIPNRPFEMKLHRRNI FEESYRRIMS VKRPDVLKARLWIFESEKGLDYGV 635  
  
 QY 410 AREWYLLCHEMLNPGYLFQYSTONIYMLQINPDSSI-NPDHLSYFHFVGRIMGLAVFH 468  
 DB 636 AREWFPLLSKEMFNPGYLFYSACNTYTLQINPNSGLCNEHDHLSYFTTIGRAGIYV 695





Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier A.C., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassaman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*.";  
 Science 287:2185-2195(2000).  
 [2]  
 SEQUENCE FROM N.A.  
 Cornell M., Evans D.A.P., Mann R., Postier M., Flaszka M., Monthathong M., Arcaivanis-Tsakonas S., Baron M.;  
 "The *Drosophila melanogaster* Suppressor of deltex gene, a regulator of the Notch receptor signalling pathway, is an E3 class ubiquitin ligase";  
 Genetics 152:0-0(1999).  
 CC -/- SIMILARITY: CONTAINS 1 C2 DOMAIN.  
 DR EMBL; AE003584; AAF51312.1; -.  
 DR EMBL; AF152865; AAD38975.1; -.  
 DR HSP; Q13526; IPIN.  
 DR FlyBase; Fgn003557; Su(dx).  
 DR GO; GO:0007219; P-N signalling pathway; IGI.  
 DR GO; GO:0008587; P-wing margin morphogenesis; IGI.  
 DR GO; GO:0008586; P-wing vein morphogenesis; IGI.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR008973; C2\_CaB.  
 DR InterPro; IPR000569; HECT domain.  
 DR InterPro; IPR001202; WW\_Rsp5\_WMP.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00632; HECT; 1.  
 DR Pfam; PF00357; WW; 4.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00119; HECT; 1.  
 DR SMART; SM00456; WW; 3.  
 DR PROSITE; PS50004; C2\_DOMAIN\_2; 1.  
 DR PROSITE; PS50237; HECT; 1.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 3.  
 DR PROSITE; PS50020; WW\_DOMAIN\_2; 2.  
 SQ SEQUENCE 949 AA; 107966 MW; 74B17A8B05AC6B6B CRC64;  
 Query Match 31.9%; Score 1239; DB 5; Length 949;  
 Best Local Similarity 32.3%; Pred. No. 1.4e-92;  
 Matches 301; Conservative 133; Mismatches 251; Indels 246; Gaps 27;  
 Qy 8 RLTVLCAKLAKKDFRLDPFAKIVDVGSGCHSTDTVKNTLDPKNOHYDLYVGKTD 67  
 Db 50 QLSVTTEASLRNGFLKFNPLVELLIISK-K-RTDLVKNLSYLPKKNSEFTVLITNST 108  
 Qy 68 ITTISVNNKHKKKQ--GAGFLGCVRL--  
 Db 109 LHFVLVDHSSFRKDALMERINLAHILQHYNGRCFELELTIDLFTVSKSDNRQTKSGEL 168  
 Qy 106 -----QRDLCKL-----  
 Db 169 VAILNGLKIDMSKLQIQPVAGQNGNPPVQAVNPSVSDAAGASCMYIGGVARMELRS 228  
 Qy 136 ---GTGSGVDRCGLLENEGTVYDSDGPRPL-----SCFMEEPAPYTDTGAA--- 181  
 Db 229 SSGNSNG--ETRSLPNGGDHRRSTCAPVWEEQQQQQSQNQQLPLRMVNGSAAVPP 285  
 Qy 182 -----ACGNC-----RFVESP----- 193  
 Db 286 TAPYPOQFPAPALAREPLTVQYALPENTQPAAYLPPAGGGAAGVPGVAGPFIEQGVGL 345  
 Qy 194 ---SDORLQAO-----RLR-----NP 207  
 Db 346 PVSQSTDPQLQTPADDEPLPAGWEIRLDQYGRYYVDHNTSTYWEKPTLPFGWEIRK 405  
 Qy 208 DVRGSL-----QTPQNRPHG-----HSP----- 226

```
Db 406 DGRGRVYVDHNRKTTWCPNPSRLMHFQHWQGRHVVSGNQRYLSYQQQQOQPTAVT 465
Qy 227 -----ELPEGVEQRTVGGQYVELHTQGVSTWHDPRIP-RDLNSVNCDELGP 273
Db 466 AQVQDDEALQPLDPGWEKKIQSNRVIFVNHKRTTQWEDPRTOGQEVSLIN---EGP 522
Qy 274 LPPGWEVSTVSGRIYFVDHNNRTTQFTDPRLLHIMNHQCOLKEPSQPLPLPSSGLEDE 333
Db 523 LPPGWEIRVTAAGERFFVDHNNRTTTFEDPR-----PGAP-----XGAGVY 564
Qy 334 ELPAQRYERDLVOKLVLRLHLSLQOPQAGHCRIVSREEIFEESSYROIIMKRPKDLKKR 393
Db 565 GVP-RAYERSFRWKLSPRY-LCQSNALPSHTKITVTQTUFDSDYHQIMLPAYELRR 622
Qy 394 LNVKFRGEGLDYGGVAREWLYLLCHEMLNPYGLFOYSTDNIMYLNQINPDSSINPDHLS 453
Db 623 LVIIIFRGEGLDYGGVAREWLYLLCHEMLNPYGLFOYSTDNIMYLNQINPDSSINPDHLS 682
Qy 454 YPHFVGRINGLAVFHGHYINGGFTVPFYKQLLGKPIQLSDLESVDPELHKSILVILENDI 513
Db 683 YPKFICGRTAMALYHGRFTYSGFTMPFYKRMNLKLTIDTIDTIDTIDTIDTIDTIDTID 742
Qy 514 TPDV-LDHTFCVHNHAFGRILQHELKPNRNVPTVENKKEYVRLVYNNRFRMGTEAFLA 572
Db 743 DCCGLELWFSVDVFLVGLIIHHLKENGKEKERTVENKEEYITLWTRMTRGIEQQTKT 802
Qy 573 LQGFNELIPOHLLKPPDOKELELIIGLDKIDLNDWKSNTLRKHCVADSNIVRWFQAV 632
Db 803 FLEGNEVPLEWLYFDEREDELICGMQDVEDQWRNTTYRHYNRNSQVWVWFQV 862
Qy 633 EYFDERRARLQFVTSRVPVLOGFKALQGSTGAAGRLFTIHLIDANTDNLKPAHCTCF 692
Db 863 RETDNEKARLQFVTSRVPVLOGFKALQGSTGAAGRLFTIHLIDANTDNLKPAHCTCF 918
Qy 693 NRIDIPPEYSEKLYEKLITAVEETCGFAVE 723
Db 919 NRIIDIPPEYSEKLYEKLITAVEETCGFAVE 949

RESULT 14
Q8WU9 PRELIMINARY; PRT; 855 AA.
ID Q8WU9 PRELIMINARY; PRT; 858 AA.
AC Q8WU9 PRELIMINARY; PRT; 858 AA.
DT 01-WAR-2002 (TREMBlrel. 20, Created)
DT 01-WAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019345.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0005512; F:ubiquitin cycle; IEA.
DR InterPro; IPR008973; C2:CaLB.
DR InterPro; IPR000569; HECT domain.
DR InterPro; IPR002349; WW_.
DR InterPro; IPR001202; WW_rsp5_WWP.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00119; HECT; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS0237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN; 1; 4.
DR PROSITE; PS0020; WW_DOMAIN; 2; 4.
```

```
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 855 AA; 98604 MW; 3CF4A66996F033EA CRC64;

Query Match 31.6%; Score 1228.5; DB 4; Length 855;
Best Local Similarity 40.4%; Pred. No. 8.7e-22; Indels 93; Gaps 19;
Matches 276; Conservative 103; Mismatches 22;

Qy 93 LSNARSL-----KDTGYQLDLCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134
Db 207 LSELSRRLQITPDSNGEQFSSLIQREPSSRLSSCVTDAVAEQGHLPSPSVAVVHTTPG 266
Qy 135 IGTGGSV-VDCRGLL-----ENEGTVYDSSGPRPLSCFWEPPAPYTDSTGAAGGN-- 186
Db 267 LPSGWEERKDAKRTYVYVNNHRTTWT-----RPMQLAEDGA-----SGSATNSNNHL 316
Qy 187 -----CRVEGPSODORLQORLNPDPVRGSL-----QTPQNRPHGHQSPE----- 227
Db 317 IEQIRPRSLSTFTVTLSPLESAKDSPVRAVKDTLSNPQSPQSPYNSPKQHKVTQ 376
Qy 228 --LPEGVEQRTVGGQYVELHTQGVSTWHDPRIP-----RDLNSVNCDELGPDPGWEV 280
Db 377 SFLPPGWEMEIAPNGRPFIDHNTKTTTWDPRKLPFVHMRSKTSLNPDLGLPLPGWEE 436
Qy 281 RSTVSGRIYFVDHNNRTTQFTDPRLLHIMNHQCOLKEPSQPLPLPSEGSLEDEELPAQY 340
Db 437 RIHLDGRTFIDHNSKITQWEDPRLQN-----PAITG-----PAVPY 473
Qy 341 ERDLVQKLVRLHLSLQOPQAGHCRIVSREEIFEESSYROIIMK-RPKDLKKRLMYKFR 399
Db 474 SREFKQYDYFRKKLKKPADIPNRFEKMLHNNIFBESYRIRMSVKRDPVLKARLWIEFE 533
Qy 400 GEGLDYGGVAREWLYLLCHEMLNPYGLFOYSTDNIMYLNQINPDSSI-NPDHLSYPHFV 458
Db 534 SEKLDYGGVAREWLYLLCHEMLNPYGLFOYSTDNIMYLNQINPDSSINPDHLSYPHFV 593
Qy 459 GRIMGLAVFHGHYINGGFTVPFYKQLLGKPIQLSDLESVDPELHKSILVILENDITPVL 518
Db 594 GRVAGLAVFHGHYINGGFTVPFYKQLLGKPIQLSDLESVDPELHKSILVILENDITPVL 652
Qy 519 HTFCVEHNAFGRILQHELKPNRNVPTVENKKEYVRLVYNNRFRMGIEAQFLAQGFN 578
Db 653 LMFCDISENFQTVQVDLKPNGSEIMVTNENKKEYIDLVIQWRFVNRVQKMAFLEGFT 712
Qy 579 ELIPCHLLKPPDOKELELIIGLDKIDLNDWKSNTLRKHCVADSN-IVRWFQAVETFE 637
Db 713 ELLPIDLIKIFDENELELLMCGLDVVDVNDVROHSIYKNGYCNHPVIQWFKAVLLMDA 772
Qy 638 ERRARLQFVTSRVPVLOGFKALQGSTGAAGRLFTIHLIDANTDNLKPAHCTCFNRDI 697
Db 773 EKIRLLQFVTGTSRVPVMPNGFAELYGSN---GPQLFTIEQW-GSPEKLPRAHTCFNRDL 828
Qy 698 PPYESYEKLYEKLITAVEETCGF 720
Db 829 PPYETFDLREKLLMAVENAQGF 851

RESULT 15
Q8BW58 PRELIMINARY; PRT; 858 AA.
ID Q8BW58 PRELIMINARY; PRT; 858 AA.
AC Q8BW58 PRELIMINARY; PRT; 858 AA.
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
```

RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC000621; AAH00621.1; -.

DR HSSP; Q13526; 1PIN.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.

DR GO; GO:0006512; P:ubiquitin cycle; IEA.

DR InterPro; IPR008973; C2 CalB.

DR InterPro; IPR000569; HECT\_domain.

DR InterPro; IPR002349; WW.

DR InterPro; IPR001202; WW\_Rep5\_WWP.

DR Pfam; PF00632; HECT; 1.

DR Pfam; PF00397; WW; 4.

DR PRINTS; PRO0403; WWDOMAIN.

DR SMART; SM00119; HECTc; 1.

DR SMART; SM00456; WW; 4.

DR PROSITE; PS0237; HECT; 1.

DR PROSITE; PS01159; WW DOMAIN 1; 4.

DR PROSITE; PS00020; WW DOMAIN 2; 4.

KW Hypothetical protein.

FT NON TER 1

SQ SEQUENCE 858 AA; 98918 MW; E2E38337F51162011 CRC64;

Query March 31.6%; Score 1228.5; DB 4; Length 858;

Best Local Similarity 40.4%; Pred.No. 8.8e-92;

Matches 276; Conservative 103; Mismatches 211; Indels 93; Gaps 19;

QY 93 LSNAISRL-----KDTGYQLDLCKLNPSD-----TDAVRGQ-----IVVSLQTRDR 134

Db 210 LSELSRLQITPDSNGEQSSLIQREPSRLRSCSVTDAVQGHLPSPSVAYVHTTPG 269

QY 135 IGTGGSV-VDCRGLL-----ENECTYVEDSGPRLSCFMEEPAPYTDSTGAAGGCGN-- 186

Db 270 LPSCWEERKQAKGRYYVNNHNRITTTWT-----RPIMQLAEDGA-----SGSATNSNNHL 319

QY 187 -----CRFVESPSQORLQAORLNRDPVGRSL-----QTPQNRPHGHQSPE----- 227

Db 320 IEPQIRPRSLSTVTLSAPLEGAKDSPVRAVKDTLSNPQSPQSPYNSPKQHKVTQ 379

QY 228 --LPEGYEQRITVGQVYFLHTQTGVSTWHDPRIP-----RDLNSVNCDELGPLPGWEV 280

Db 380 SFLPPGWMEMRIAPNGRPFFIDHNTKTTTWEDPRLKFPVHMRKSTLSLNPDLGPLPGWEE 439

QY 281 RSTVSGRIYFYVDHNRRTQFTDRELHIMHQCQLKEPSQPLPLPSEGSLEDEELPAQRY 340

Db 440 RIHLDGRTFYIDHNSKITQWEDPRLQN-----PATG-----PAVPY 476

QY 341 ERDLVQKLKVLRLHLSLQOQAGHCRIEVSREEIFEESYRQIMKM-RPKDLKKRLMYKFR 399

Db 477 SREFPKQYDYPRKKLKPADIPNRFEMKLRHNNIFEESYRIRMSVKRPDVLKARLWIEPE 536

QY 400 GEEGLDYGGVAREWLYLLCHEMLNPYYGLFOYSTDNITVMLQINPDSSI-NPDHLSYHFV 458

Db 537 SEKGLDYGGVAREWFLLSKEMFNPYYGLPEYSATDNYTLQINPNSGLCNEDHLSYFTFI 596

QY 459 GRIMGLAVFHGHYINGGFTVPFFYKQLLQKPIQLSDLESVDPELHKSILWILENDITPVLD 518

Db 597 GRVAGLAVFHGKLDGDFIRFFYQWMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 655

QY 519 HTFCVENHAFGRILQHELKPNGRNVPVTENKKEYRLYVNNWRPMRGIEAQFLAQKGFN 578

Db 656 LMFCDIEENFQTYQVDLKPNGSIVMTNENKREYIDLVIQWRVNRVQKQNAFLLEGFT 715

QY 579 ELIPQHLKPPDOKELELIIIGLLDKILDNDWKNTRLKHCVDASN-IVRWFWQAVETFE 637

Db 716 ELLPIDLIKIFDELELELMCGLDGVDVNDWRQHSIYKNGYCFNHPVIOFWKAVLLMDA 775

QY 638 ERRARLLQFVTGTRVPLQGFKALQSGTGAAGPRLFTIHLIDANTDNLPAKHTCFNRIDI 697

Db 776 EKIRLLQFVTGTRVPMNGFAELYGSN---GPQLFTIEQW-GSPEKLPRAHTCFNRLLD 831

QY 698 PPYESYKLYEKLITAYEETCGF 720

Db 832 PPYETFDLREKLLMAVENAQGF 854

Search completed: September 21, 2004, 07:47:22  
Job time : 96.9789 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:47:37 ; Search time 91.9109 Seconds  
(without alignments)  
2526.134 Million cell updates/sec

Title: US-10-009-945-2

Perfect score: 3884

Sequence: 1 GSSIKIRLTVLCAKNAK...EKLVEKLTAVETCGFAVE 723

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3878	99.8	722	14	US-10-097-534-14
2	3001	77.3	748	14	US-10-021-660-81
3	2996	77.1	804	12	US-10-220-120-407
4	2913.5	75.0	735	14	US-10-313-985-2
5	2236	57.6	514	11	US-09-764-875-819
6	1471	37.9	766	14	US-10-313-955-4
7	1436.5	37.0	832	14	US-10-032-585-7296
8	1345	34.6	869	14	US-10-128-714-8162
9	1259	32.4	927	14	US-10-097-534-15
10	1249.5	32.2	911	14	US-10-205-823-279
11	1245.5	32.1	739	12	US-10-183-936A-89
12	1245.5	32.1	739	14	US-10-097-534-10
13	1245.5	32.1	739	15	US-10-374-979-89
14	1245.5	32.1	732	10	US-09-919-039-235
15	1228.5	31.6	834	14	US-10-313-955-6

16	1228.5	31.6	995	14	US-10-097-534-9	Sequence 9, Appl
17	1228.5	31.6	995	14	US-10-205-823-275	Sequence 275, App
18	1227	31.6	725	14	US-10-185-050-126	Sequence 126, App
19	1226.5	31.6	759	14	US-10-128-714-3162	Sequence 3162, Ap
20	1226	31.6	854	14	US-10-205-823-277	Sequence 277, App
21	1225	31.5	870	14	US-10-097-534-12	Sequence 12, Appl
22	1225	31.5	906	14	US-10-185-050-48	Sequence 48, Appl
23	1223	31.5	854	14	US-10-287-218-3	Sequence 3, Appl
24	1223	31.5	854	16	US-10-474-291-3	Sequence 3, Appl
25	1212	31.2	474	10	US-09-774-639-371	Sequence 371, App
26	1212	31.2	474	10	US-09-969-730-249	Sequence 249, App
27	1212	31.2	474	16	US-10-621-363-249	Sequence 249, App
28	1195	30.8	898	12	US-10-188-186-114	Sequence 114, App
29	1047	27.0	683	14	US-10-185-050-46	Sequence 46, Appl
30	1047	27.0	684	14	US-10-097-534-11	Sequence 11, Appl
31	1023	26.3	380	14	US-10-307-956-32	Sequence 32, Appl
32	1012	26.1	375	14	US-10-307-956-31	Sequence 31, Appl
33	1002	25.8	1562	14	US-10-097-534-13	Sequence 13, Appl
34	836.5	21.5	1094	14	US-10-043-487-300	Sequence 300, App
35	836.5	21.5	1488	14	US-10-043-487-285	Sequence 285, App
36	836.5	21.5	2011	16	US-10-408-765A-767	Sequence 767, App
37	821.5	21.2	733	14	US-10-097-534-16	Sequence 16, Appl
38	807.5	20.8	1104	12	US-10-424-599-145002	Sequence 145002,
39	800	20.6	1843	12	US-10-424-599-145004	Sequence 145004,
40	776.5	20.0	3647	16	US-10-437-963-119793	Sequence 119793,
41	754.5	19.4	973	16	US-10-437-963-178922	Sequence 178922,
42	724	18.6	277	9	US-09-925-300-1628	Sequence 1628, Ap
43	720.5	18.6	358	14	US-10-268-036-5	Sequence 5, Appl
44	720.5	18.6	358	16	US-10-391-364-33	Sequence 33, Appl
45	691	17.8	157	12	US-10-211-462-107	Sequence 107, App

ALIGNMENTS

RESULT 1

US-10-097-534-14  
; Sequence 14, Application US/10097534  
; Publication No. US20030049607A1  
; GENERAL INFORMATION:  
; APPLICANT: GREENER, TSVIKA  
; APPLICANT: MOSKOWITZ, HAIM  
; APPLICANT: REISS, YUVAL  
; APPLICANT: ALROY, IRIS  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL  
; TITLE OF INVENTION: NATURATION  
; FILE REFERENCE: PLV-001.01  
; CURRENT APPLICATION NUMBER: US/10/097,534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/275,224  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/308,958  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/340,170  
; PRIOR FILING DATE: 2001-12-07  
; NUMBER OF SEQ ID NOS: 71  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 722  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-534-14

Query Match 99.8%; Score 3878; DB 14; Length 722;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GSSIKIRLTVLCAKNAK...EKLVEKLTAVETCGFAVE 723

Db 1 GSSIKIRLTVLCAKNAK...EKLVEKLTAVETCGFAVE 60

Qy 62 VKGTDSTIISVNNKHKTKGAGFLGCVRLLSNAISRLKDTGYCRDLCKLNESDIDAV 121

Db 61 VGKTDITISVNNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQRLDCLKNPSTDVA 120  
Qy 122 RGOIVSLQTDRIQTGSGVVDRCGLLENEGTVYDSCGPPLSCFEEEPAPYTDSTGAA 181  
Db 121 RGOIVSLQTDRIQTGSGVVDRCGLLENEGTVYDSCGPPLSCFEEEPAPYTDSTGAA 180  
Qy 182 AGGNCRFVESPSQDRLQARLNPDPVRSGLQTPQNRPHGHSQPELPEGYEQRTTVQGG 241  
Db 181 AGGNCRFVESPSQDRLQARLNPDPVRSGLQTPQNRPHGHSQPELPEGYEQRTTVQGG 240  
Qy 242 VYFLHTQTVSTWHPDPRIPRINSVNCDELGCPBPCHVEVSTSGRIYFVDHNNRTTQFT 301  
Db 241 VYFLHTQTVSTWHPDPRIPRINSVNCDELGCPBPCHVEVSTSGRIYFVDHNNRTTQFT 300  
Qy 302 DPLHHMNHQOLKPEPSQPLPSEGSLEDEELPAQRYERDLVQKLKVRHLSLQOPQ 361  
Db 301 DPLHHMNHQOLKPEPSQPLPSEGSLEDEELPAQRYERDLVQKLKVRHLSLQOPQ 360  
Qy 362 AGHCRIVSRBEIPEESYRQIMKWRPKDLKRLMKVRGEBGLDYGAVREWLILLCHEM 421  
Db 361 AGHCRIVSRBEIPEESYRQIMKWRPKDLKRLMKVRGEBGLDYGAVREWLILLCHEM 420  
Qy 422 LNPYGLFQYSTDNLYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHHYINGGTVFFY 481  
Db 421 LNPYGLFQYSTDNLYMLQINPDSSINPDHLSYFHFVGRIMGLAVFHHYINGGTVFFY 480  
Qy 482 KOLGKPIQLSDLESVDPELHKSLLWILENDITPVLDTFCVEHNAFGRILQHELKPNR 541  
Db 481 KOLGKPIQLSDLESVDPELHKSLLWILENDITPVLDTFCVEHNAFGRILQHELKPNR 540  
Qy 542 NVPVTEENKKEVRLVNNRFRWGIEAQLALQGENELIPQHLKPPDOKLELIIIGGL 601  
Db 541 NVPVTEENKKEVRLVNNRFRWGIEAQLALQGENELIPQHLKPPDOKLELIIIGGL 600  
Qy 602 DKIDLNDKSNTRLKHCVADSNIVRWFQAVETDEERRALLQFVTGSTRVPLQGFAL 661  
Db 601 DKIDLNDKSNTRLKHCVADSNIVRWFQAVETDEERRALLQFVTGSTRVPLQGFAL 660  
Qy 662 QGSTGAAGPRLFTHLIDANTDNLPAKHTCFNRDIPPESEYKLYEKLITAVEETCGFA 721  
Db 661 QGSTGAAGPRLFTHLIDANTDNLPAKHTCFNRDIPPESEYKLYEKLITAVEETCGFA 720  
Qy 722 VE 723  
Db 721 VE 722

RESULT 2  
US-10-021-660-81  
; Sequence 81, Application US/10021660  
; Publication No. US20030152926A1  
; GENERAL INFORMATION:  
; APPLICANT: Murray, Richard  
; APPLICANT: Glynn, Richard  
; APPLICANT: Watson, Susan R.  
; TITLE OF INVENTION: EOS Biotechnology, Inc.  
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis  
; TITLE OF INVENTION: Modulators  
; FILE REFERENCE: 018501-0007110US  
; CURRENT APPLICATION NUMBER: US/10/021,660  
; CURRENT FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: US/09/784,356  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: US 09/637,977  
; PRIOR FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: FastSeq for Windows Version 3.0.  
; SEQ ID NO 81  
; LENGTH: 748  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-021-660-81

Query Match 77.3%; Score 3001; DB 14; Length 748;  
Best Local Similarity 74.6%; Pred. No. 5,5e-266;  
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

Qy 5 IKIRLTVICAKMLAKKDFRLLPDPFAKIVVDGSGOCHSTDTVNTLDPKNQHVLDLYGK 64  
Db 12 VKLRLTVICAKMLVKKDFRLLPDPFAKIVVDGSGOCHSTDTVNTLDPKNQHVLDLYGK 71  
Qy 65 TDSITISVNNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQRLDCLKNPSTDVAVRGQ 124  
Db 72 SDSVITISVNNHKKIHKQAGFLGCVRLLSNAISRLKDTGYQRLDCLKNPSTDVAVRGQ 131  
Qy 125 IVVSLQTDRIQTGSGVVDRCGLLEN-----GTYY----- 155  
Db 132 IVVSLQTDRIQTGSGVVDRCGLLEN-----GTYY----- 191  
Qy 156 --EDSGRPLSCFEEEPAPYTDSTGAAAGGNCRFVESPSQDRLQARLNPDPVRSGL 213  
Db 192 AGEYSRPGPLSCFEEEPAPYTDSTGAAAGGNCRFVESPSQDRLQARLNPDPVRSGL 242  
Qy 214 QTPQNRPHGHSQPELPEGYEQRTTVQGGYFLHTQTVSTWHPDPRIPRINSVNCDELGCP 273  
Db 243 ---SRTHLHTPDLPEGYEQRTTVQGGYFLHTQTVSTWHPDPRIPRINSVNCDELGCP 298  
Qy 274 LPPGHEVSTSGRIYFVDHNNRTTQFTDPR-----LHIMNHQOLKPEPSQPLPSEGS 329  
Db 299 LPPGHEVSTSGRIYFVDHNNRTTQFTDPR-----LHIMNHQOLKPEPSQPLPSEGS 354  
Qy 330 L---EDELPAQRYERDLVQKLKVRHLSLQOPQAGHCRIVSRBEIPEESYRQIMKWR 386  
Db 355 LCPDDETECLTVPRYKEDLVQKLKVRHLSLQOPQAGHCRIVSRBEIPEESYRQIMKWR 414  
Qy 387 PKDLKRLMKVRGEBGLDYGAVREWLILLCHEMNLNYPYGLFQYSTDNLYMLQINPDSS 446  
Db 415 PKDLKRLMKVRGEBGLDYGAVREWLILLCHEMNLNYPYGLFQYSTDNLYMLQINPDSS 474  
Qy 447 INPDHLSYFHFVGRIMGLAVFHHYINGGTVFFYKQLGKPIQLSDLESVDPELHKSLLV 506  
Db 475 VNPVTEENKKEVRLVNNRFRWGIEAQLALQGENELIPQHLKPPDOKLELIIIGGL 534  
Qy 507 WILENDITPVLDTFCVEHNAFGRILQHELKPNRNPVTEENKKEVRLVNNRFRWGIE 566  
Db 535 WILENDITPVLDTFCVEHNAFGRILQHELKPNRNPVTEENKKEVRLVNNRFRWGIE 594  
Qy 567 EAQFLALQGFNEVLPQHLKTFDEKELELIIIGGLKIDVNDKVNTRLKHCVADSNIVR 626  
Db 595 EAQFLALQGFNEVLPQHLKTFDEKELELIIIGGLKIDVNDKVNTRLKHCVADSNIVR 654  
Qy 627 WFWQAVETDEERRARLLQFVTGSTRVPLQGFALQGSTGAAGPRLFTHLIDANTDNL 686  
Db 655 WFWQAVETDEERRARLLQFVTGSTRVPLQGFALQGSTGAAGPRLFTHLIDANTDNL 711  
Qy 687 KAHTCFNRDIPPESEYKLYEKLITAVEETCGFAVE 723  
Db 712 KAHTCFNRDIPPESEYKLYEKLITAVEETCGFAVE 748

RESULT 3  
US-10-220-120-407  
; Sequence 407, Application US/10220120  
; Publication No. US20040048253A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: PANZER, Scott R.  
; APPLICANT: SPIRO, Peter A.  
; APPLICANT: BANVILLE, Steven C.  
; APPLICANT: SHAH, Purvi  
; APPLICANT: CHALUP, Michael S.  
; APPLICANT: CHEN, Simon C.  
; APPLICANT: CHEN, Alice  
; APPLICANT: D'SA, Steven A.  
; APPLICANT: AMSHEY, Stefan

APPLICANT: DAHL, Christopher R.  
APPLICANT: DAM, Tam C.  
APPLICANT: DANIELS, Susan E.  
APPLICANT: DUFOUR, Gerard E.  
APPLICANT: FLORES, Vincent  
APPLICANT: FONG, Willy T.  
APPLICANT: GREENAWALT, Lila B.  
APPLICANT: HILLMAN, Jennifer L.  
APPLICANT: JONES, Anissa L.  
APPLICANT: LIU, Tommy F.  
APPLICANT: ROSEBERRY, Ann M.  
APPLICANT: ROSEN, Bruce H.  
APPLICANT: RUSSO, Frank D.  
APPLICANT: STOCKDREHER, Theresa K.  
APPLICANT: DAFFO, Abel  
APPLICANT: WRIGHT, Rachel J.  
APPLICANT: YAP, Pierre E.  
APPLICANT: YU, Jimmy Y.  
APPLICANT: BRADLEY, Diana L.  
APPLICANT: BRATCHER, Shawn R.  
APPLICANT: CHEN, Wensheng  
APPLICANT: COHEN, Howard J.  
APPLICANT: HODGSON, David M.  
APPLICANT: LINCOLN, Stephen E.  
APPLICANT: JACKSON, Stuart  
TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: PT-1113 PCT  
CURRENT APPLICATION NUMBER: US/10/220,120  
CURRENT FILING DATE: 2002-08-26  
PRIORITY APPLICATION NUMBER: 60/184,777; 60/184,797; 60/184,698; 60/184,770; 60/184,774;  
60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;  
60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;  
60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;  
60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;  
60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;  
60/205,324; 60/205,286  
PRIORITY FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;  
2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-05-17;  
2000-05-17; 2000-05-12; 2000-05-16; 2000-05-16; 2000-05-15;  
2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;  
2000-05-17; 2000-05-17  
NUMBER OF SEQ ID NOS: 422  
SOFTWARE: PERL Program  
SEQ ID NO 407  
LENGTH: 804  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20040048253A1 LG:132147.3.orf3:2000FEB18  
US-10-220-120-407

Query Match 77.1%; Score 2996; DB 12; Length 804;  
Best Local Similarity 74.5%; Pred. No. 1.8e-265;  
Matches 564; Conservative 63; Mismatches 72; Indels 58; Gaps 8;  
5 IKRLTVLCAKNAUKDFRLDPFPAKIVVDGSGQCHSTDTVNLTDPKWNQHYDLYGK 64  
68 VKLHLTVLCAKNAUKDFRLDPFPAKIVVDGSGQCHSTDTVNLTDPKWNQHYDLYGK 127  
65 TDSITISVNNKHKKQAGFLGCVRLLSNATSRKLDGTGYRLDCKLNPSDPAVRGQ 124  
128 SDSVTISVNNKHKKQAGFLGCVRLLSNATSRKLDGTGYRLDCKLNPSDPAVRGQ 187  
125 IVVLSQDRIGTGSGVDCRGLLENE-----GTVY----- 155  
188 IVVLSQDRIGTGSGVDCRGLLENE-----GTVY----- 155  
156 --EDSGPRLSCMEEPAPYDSTGAAAGGNCRFVSPDQRLQAOQLANEDVRGSL 213  
248 ASYSSPGRPLSCFVDENTPISGNTGATCG-----QSDPRLAERRVRSQRHNYM 298

214 QTPONRPHGHOSPELPEGYEORTTVOGOVYFLHTQTGVSTWHDPRIPRDLNSVNCDELGP 273  
299 -----SRTHLHTPPDLPEGYEORTTVOGOVYFLHTQTGVSTWHDPRIPRDLNSVNCDELGP 354  
274 LPPGWEVRSVSGRIYFVDHNNRTTQFTDPR-----LHHINWOCOLKEPSPQPLPSEGS 329  
355 LPPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLELVLRNQLKDDQOQVW-----S 410  
330 L---EDELPAQRYERDLVOKLVLRHLSLOQFAGHCHRIEVSREIPEESYRQIMKMR 386  
411 LCPDDTECLTVPRYKRDVLVOKLVLRHLSLOQFAGHCHRIEVSREIPEESYRQIMKMR 470  
387 PKDLKXELMVKFRGEEGLDYGVAREWLYLLCHEMLNPPYGLFYQYSTDNLYMLQINPDS 446  
471 PKDLKXELMVKFRGEEGLDYGVAREWLYLLCHEMLNPPYGLFYQYSTDNLYMLQINPDS 530  
447 INPDHLSYFHFVGRIMGLAVFHHYNGFTVFYKOLLGKPIQLSLESVDPELHKSIV 506  
531 VNPHELSYFHFVGRIMGLAVFHHYNGFTVFYKOLLGKPIQLSLESVDPELHKSIV 590  
507 WILENDITPVLDTFCVEHNAFGRILQHELKPKGRNVPVTEENKKEYVRLYVNRVFRGI 566  
591 WILENDITPVLDTFCVEHNAFGRILQHELKPKGRNVPVTEENKKEYVRLYVNRVFRGI 650  
567 BAQFLALQKGFNELIPQHLKPFDOKELELIIIGLDKIDLNDKSNTRLKHCVADSIVR 626  
651 BAQFLALQKGFNEVIPHLLKTFDEKELELIIIGLDKIDLNDKSNTRLKHCVADSIVR 710  
627 WFOAVTFTEERRARLLQFVTGSTRVPLQFGKALQSGTAGAPRFTIHLIDANTDNL 686  
711 WFKAVTFTEERRARLLQFVTGSTRVPLQFGKALQSGTAGAPRFTIHLIDANTDNL 767  
687 KAHTCFNRIDIPPPVSEYKELYKLLTAVETCGFAVE 723  
768 KAHTCFNRIDIPPPVSEYKELYKLLTAVETCGFAVE 804

RESULT 4  
US-10-313-955-2  
; Sequence 2, Application US/10313955  
; Publication No. US20030199036A1  
; GENERAL INFORMATION:  
; APPLICANT: Beach, David H.  
; Nefsky, Bradley  
; Nefsky, Bradley  
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/10/313,955  
; FILING DATE: 05-Dec-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/392,163  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 08/539,205  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: CSV-005.01  
; TELECOMMUNICATION INFORMATION:

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 819
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-875-819

Query Match      57.6%; Score 2236; DB 11; Length 514;
Best Local Similarity 80.7%; Pred. No. 5.6e-196;
Matches 414; Conservative 43; Mismatches 42; Indels 14; Gaps 4;

QY 218 NRPHGQSPPELPEGEYQRTTVOGVYFLHTQTGVSTWHDPRIPRDLNSVNCDELGLPPGP 277
DQ 9 SRTLHTPPDLPEGEYQRTTVOGVYFLHTQTGVSTWHDPRIPRDLNSVNCDELGLPPGP 68
QY 278 WEVRSTVSGRIYFVDHNNRTTQTDPR---LHHIMNHQCOLKEPSPQLPLPSGSL--- 330
DQ 69 WEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVNLNQKDKQQQQVV---SLCPD 124
QY 331 EDEELPAQRYERDLVQKLVRHELSQLQPOAGHCRIEVSREEIFEESYRQIMKMRPKDL 390
DQ 125 DTECLTVPRYKEDLVQKLVRHELSQLQPOAGHCRIEVSREEIFEESYRQIMKMRPKDL 184
QY 391 KRLMKVRGEBGLDYGAVAREWLYLLCHEMLNPPYGLFOYSTDNIMYLNQINPDSSINPD 450
DQ 185 WKRLMKVRGEBGLDYGAVAREWLYLLCHEMLNPPYGLFOYSTDNIMYLNQINPDSSINPD 244
QY 451 HUSYFHFVGRIMGLAVPHGHYINGGFTVPFYKQLGKPIQLSDLESVDPELHKSILVLE 510
DQ 245 HUSYFHFVGRIMGLAVPHGHYIDGGFTLPFYKQLGKSIITLDDMELVDPDLNSILVLE 304
QY 511 NDTIPVLDTFCVEHNAFGRILQHELKPNRNPVTEENKKEYVRLYVNWRFMRGIZAQF 570
DQ 305 NDTIGVLDHTFCVEHNAFGRILQHELKPNRNPVTEENKKEYVRLYVNWRFMRGIZAQF 364
QY 571 LALQKGFNEVLPQHLLKTFDEKLELIIICGLKIDVNDKNTLKHCTPDSNVKFWK 424
DQ 365 LALQKGFNEVLPQHLLKTFDEKLELIIICGLKIDVNDKNTLKHCTPDSNVKFWK 424
QY 631 AVETFEDEERRARLLQFVTGSTRVPLQGFALQSGSTGAAGPRLFTIHLIDANTNLPKHA 690
DQ 425 AVEFDEERRARLLQFVTGSTRVPLQGFALQSGSTGAAGPRLFTIHLIDANTNLPKHA 481
QY 691 CNRNDIPPEYSEYKLYEKLITAVEETCGFAVE 723
DQ 482 CNRNDIPPEYSEYKLYEKLITAVEETCGFAVE 514

RESULT 6
US-10-313-955-4
; Sequence 4, Application US/10313955
; Publication No. US20030199036A1
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; Nefsky, Bradley
; Caligiuri, Maureen
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 819
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-875-819

Query Match      75.0%; Score 2913.5; DB 14; Length 735;
Best Local Similarity 72.9%; Pred. No. 5.8e-258;
Matches 552; Conservative 63; Mismatches 71; Indels 71; Gaps 9;

QY 5 IKRLITVLCANLAKKOFFRPLDPDFAKIVDGSQCHSTDTVKNLTPKKNQHYDLVYVK 64
DQ 12 VKRLITVLCANLAKKOFFRPLDPDFAKIVDGSQCHSTDTVKNLTPKKNQHYDLVYVK 58
QY 65 TDSITISVNNHKKIKKQAGAGLGCVRLLSNAISRLKDTGYQRLDCKLNPSTDAVRGQ 124
DQ 59 SDSVITISVNNHKKIKKQAGAGLGCVRLLSNAISRLKDTGYQRLDCKLNPSTDAVRGQ 118
QY 125 IVVSLQTRDRITGSGSVVDCGLLENE-----GTVY----- 155
DQ 119 IVVSLQTRDRITGSGSVVDCGLLENE-----GTVY----- 178
QY 156 --EDSGPGRPLSCWEEBAPYTDSTGAAAGGNCRFVESPSQDORLQAORLRNPDRVGL 213
DQ 179 ASEYSSPGRPLSCWEEBAPYTDSTGAAAGGNCRFVESPSQDORLQAORLRNPDRVGL 229
QY 214 QTPQNRPHGOSPELPEGEYQRTTVOGVYFLHTQTGVSTWHDPRIPRDLNSVNCDELGP 273
DQ 230 ---SRTLHTPPDLPEGEYQRTTVOGVYFLHTQTGVSTWHDPRIPRDLNSVNCDELGP 285
QY 274 LPPGWEVRSTVSGRIYFVDHNNRTTQTDPR---LHHIMNHQCOLKEPSPQLPLPSGSL 329
DQ 286 LPPGWEVRSTVSGRIYFVDHNNRTTQTDPRLSANLHLVNLNQKDKQQQQVV---S 341
QY 330 L---EDELPAQRYERDLVQKLVRHELSQLQPOAGHCRIEVSREEIFEESYRQIMKMR 386
DQ 342 LCPDDTECLTVPRYKEDLVQKLVRHELSQLQPOAGHCRIEVSREEIFEESYRQIMKMR 401
QY 387 PKDLKRLMKVRGEBGLDYGAVAREWLYLLCHEMLNPPYGLFOYSTDNIMYLNQINPDSS 446
DQ 402 PKDLKRLMKVRGEBGLDYGAVAREWLYLLCHEMLNPPYGLFOYSTDNIMYLNQINPDSS 461
QY 447 INPDHLSYFHFVGRIMGLAVPHGHYINGGFTVPFYKQLGKPIQLSDLESVDPELHKSILV 506
DQ 462 VNPHELVSYFHFVGRIMGLAVPHGHYIDGGFTLPFYKQLGKSIITLDDMELVDPDLNSILV 521
QY 507 WILENDITPVLDHTFCVEHNAFGRILQHELKPNRNPVTEENKKEYVRLYVNWRFMRGI 566
DQ 522 WILENDITPVLDHTFCVEHNAFGRILQHELKPNRNPVTEENKKEYVRLYVNWRFMRGI 581
QY 567 BAQFLALQKGFNEVLPQHLLKTFDEKLELIIICGLKIDVNDKNTLKHCTPDSNVK 626
DQ 582 BAQFLALQKGFNEVLPQHLLKTFDEKLELIIICGLKIDVNDKNTLKHCTPDSNVK 641
QY 627 WFWQAVETFEDEERRARLLQFVTGSTRVPLQGFALQSGSTGAAGPRLFTIHLIDANTNLP 686
DQ 642 WFWQAVETFEDEERRARLLQFVTGSTRVPLQGFALQSGSTGAAGPRLFTIHLIDANTNLP 698
QY 687 KAHTCFNRNDIPPEYSEYKLYEKLITAVEETCGFAVE 723
DQ 699 KAHTCFNRNDIPPEYSEYKLYEKLITAVEETCGFAVE 735

RESULT 5
US-09-764-875-819
; Sequence 819, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
```





```

; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 927
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-10-097-534-15

Query Match
Best Local Similarity 32.4%; Score 1259; DB 14; Length 927;
Best Local Similarity 33.6%; Pred. No. 8.1e-106;
Matches 307; Conservative 129; Mismatches 239; Indels 240; Gaps 28;

QY 7 IRLTVLCANLAKKDPFRLPDPFAKIV---VDGSGQCHSDTDTVNTLDPKWN-----55
DB 48 VRVRVIAGILGAKKDILOGASDPIVATVLDPMNGVLTSTVQTKTKSLNPKWNEILFRV 107
QY 56 ---QH-----YDLVVGKTD---SITISVWN-----HKIKHKQG 83
DB 108 HPQOHELLFEVDENELTRDDFLGQVDVPLPLPTENRPLRPYTFKDFVLPFASHSRV 167
QY 84 AGFLGCVRLLSNAIRLKT-----GYORLD-----LCKL-----113
DB 168 KGYL-----RLKNTYLPKTSGBEDNAEQAELEFGVWVLDQPDAAACHLQOQOE 216
QY 114 -----NPSD---TDAVRGQIVSLQ-----TRDR 134
DB 217 PSLPLPGWEERQDILGRYYVNHESRRTCWKKEPTFDNLTDAGNQLQQAQAFTRRQ 276
QY 135 IGTGGSVDPGRLLN-----EGTVYEDSG-PGRPLSCFMEEFPAYTD-----176
DB 277 ISEBESVDNQESSENWEIREDATMYSSQAFFPPPPSSNLDVPTHLAEELNARLTIFG 336
QY 177 ---STGAAGGNCR-----FVESPS-----QDORLQAQRL-- 204
DB 337 NSAVOPASSNHSSRSGSLQAYTFEEQPTLPVLLPTSSGLPPGWEKQERGSYYVDH 396
QY 205 -----RNPDRVGSLOTPO-----NRPHGHSPELPEGYEQ 234
DB 397 NSRTTWTXPTQATVETISQLTSSOSSAGSQASQSDSQGVQTPSEIEGFLPKGWEV 456
QY 235 RTTVQGVFLHTQGVSTWHPDR---IPRDLNSV-----NCDELGPLPGWEVRVTSGR 288
DB 457 RHAPNGRPFFIDHNTKTTTWEDPRLKIPAHLRGKTSLTDNSDLGLPLPGWEERTHTDGR 516
QY 289 YFVDHNNRTQTDPRLHHMHQOLKEPSQPLPSEGSLEDEELPAQRYERDLVOKL 348
DB 517 FYINNIKTQWEDPLENV-----AIGTAPV-----YSDDYKRY 553
QY 349 KVLRLHLSLQOPQAGHCRIVRESREIFESYRQIMMMPKD-LKRLMVFRGEGLDYG 407
DB 554 EPPRRKLKQNDIPNKFEMKLRATVLEDSYRIMGVRRADFLKARLWIEFDGEGLDYG 613
QY 408 GVAREWLYLLCHEMLNPPYGLFQYSDNTYMLQINPDSSI-NPDLHSLYELHVGIMGLAV 466
DB 614 GVAREWFLISKEWFPYGLFEYSATONYTLQINPNSGLCNEHLSTYFKFTIGVAGMAV 673
QY 467 FHGYINGGFTVPFYKQLLGKPIQLSDLESVDPELHLSLVMLENDITPVLDDHTFCVEHN 526
DB 674 YHGKLLDGFIFRPFYKQWLHKPI-LLHDMESVDSEYNSLRWILENDPTE-LDLRFLIDEE 732
QY 527 AGRILLOHELKPNRNVPTENKKEYVRLYVNRPFMGIEAQFLAQKGNELIPQHL 586
DB 733 LFGQTHCHELKNIGSSIVVNTNKNKEYLYLQWRFVNRIQKMAAFKEGFFELIPQDL 792
QY 587 KPFDQKELELITGGDLKDILNDWKNTRLKH-CVADSNIRVFWFOAVETFEERRARLLQ 645
DB 793 KIFENELELLMCLGDVDVNDWREHTKYNGYSANHOVIQFWKAVLMDSKRIELLO 852
QY 646 FVTGSTRVLPQGFKALQGTGAAGRLFTIHLIDANTONLPAKTCNFRNIDIPYESYEK 705
DB 853 FVTGSTRVPMWNGFAELIYGN---GQOSTVEOW-GTPEKLPRAHTCNRLDLPYEGFEE 908

; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 927
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-10-097-534-15

Query Match
Best Local Similarity 32.4%; Score 1249.5; DB 14; Length 911;
Best Local Similarity 34.5%; Pred. No. 5.9e-105;
Matches 303; Conservative 120; Mismatches 243; Indels 211; Gaps 24;

QY 17 LAKXDFFLP---DPFAKI---VVDGSGQCH---STDTVNTLDPKWNQHYDLYVGKTD-S 67
DB 69 LPTSDFLPLGICDPYVKLSLYVADENRELAVQTKTKTLNPKWNEBFFRVNPSNHR 128
QY 68 ITISVWNHKKHKQAGAGFLGCVPLLSNAI-----SRDKOT 103
DB 129 LLFEVFDNRTRDD---FLGQVDVPLSHLPTEDPTMERPYTPKDFLLRPRSHKSRV-- 183
QY 104 GYORLDLCKLNPSTDVAVRQIVVSLQTRDRIGTGGSVVDRCG-----146
DB 184 GFLRLKWAYMPKNG-----GQDENSDQRDDMEHGWEVVDNSDSAQHQBELPPPLPG 238
QY 147 ---LLENBGTVY-----ED 157
DB 239 WEEKVDNLGRTYVYNNHNRRTQWHRPSLMDVSDSNINRQINQIEAAHFRFRSRHSED 298
QY 158 SGP-----GRPLSCFMEEPAFYTDSTGAA-----AGGCNCRFVESPQD-----196
DB 239 LEPSPSEGDPVEPWEITISEVNIAGDSLGLALPPPPASPGSTSPQELSEELSRLOIT 358
QY 197 -----QLQAQRLN-----PDVRGSLQTPON 218
DB 359 PDSNGEOPSSLIQEPSSRLRSCSVTDVAEQHQLPFGAKDSFVRRAVKDTLSNPQSPQ 418

; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 927
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-10-097-534-15

Query Match
Best Local Similarity 32.2%; Score 1249.5; DB 14; Length 911;
Best Local Similarity 34.5%; Pred. No. 5.9e-105;
Matches 303; Conservative 120; Mismatches 243; Indels 211; Gaps 24;

QY 17 LAKXDFFLP---DPFAKI---VVDGSGQCH---STDTVNTLDPKWNQHYDLYVGKTD-S 67
DB 69 LPTSDFLPLGICDPYVKLSLYVADENRELAVQTKTKTLNPKWNEBFFRVNPSNHR 128
QY 68 ITISVWNHKKHKQAGAGFLGCVPLLSNAI-----SRDKOT 103
DB 129 LLFEVFDNRTRDD---FLGQVDVPLSHLPTEDPTMERPYTPKDFLLRPRSHKSRV-- 183
QY 104 GYORLDLCKLNPSTDVAVRQIVVSLQTRDRIGTGGSVVDRCG-----146
DB 184 GFLRLKWAYMPKNG-----GQDENSDQRDDMEHGWEVVDNSDSAQHQBELPPPLPG 238
QY 147 ---LLENBGTVY-----ED 157
DB 239 WEEKVDNLGRTYVYNNHNRRTQWHRPSLMDVSDSNINRQINQIEAAHFRFRSRHSED 298
QY 158 SGP-----GRPLSCFMEEPAFYTDSTGAA-----AGGCNCRFVESPQD-----196
DB 239 LEPSPSEGDPVEPWEITISEVNIAGDSLGLALPPPPASPGSTSPQELSEELSRLOIT 358
QY 197 -----QLQAQRLN-----PDVRGSLQTPON 218
DB 359 PDSNGEOPSSLIQEPSSRLRSCSVTDVAEQHQLPFGAKDSFVRRAVKDTLSNPQSPQ 418

```

```
QY 219 RPHGOSPE-----LPEGYEORTTVQGVYFLHTGTGVTWHDPRIP-----RDLSNV 266
D 109 SFASATSEDSGSGSLPPTNTNTNTSEGTSLIPL-----TISGGSGP-RPLNPV 159
QY 168 MEPEPAPYTDSTGAAAGGNCRFVESPSQDQLQOARLNPDPVRSGL-----QTPQNR 219
D 160 TQAPLP-----PGWEQVR-----DQGRVYVVDHVEXRTTMDR 192
QY 220 RHGHOSPELPEGEORTTVQGVYFLHTGTGVTWHDPRIP----- 260
D 193 PE-----PLPGWEARVDNMGRIYVDFHTTTRTTWQRTLESVRNVEQWQORQLOQAM 247
QY 261 -----RDL-----NSVNCDELPLPGWEVRSTVGRIYVVDHNNRTTQTDPR 305
D 248 QCFNQRFYGNQDLFATQSKEFDPLGPLPGWEKRTDSNGRVFVNHNTRITQWEDPR- 306
QY 306 HHMMHQCOLKEPSQPLP-----PSGSLDEBELPAQR 339
D 307 -----SQGLNE--KPLPEGWEMRTVGPYFVDHNNRTTYIDRTGKALDNGPOIA 359
QY 340 YERDLVQKLKVLK-----HELSQLQOPQAGHCRIEVSREEIFESYQIMKVRKDKRLMV 396
D 360 YVRDFKAKVQYFRFCQQLAMPQ-----HIKITVTRKTLTFEDSFQOIMGFSQDLRRRLMV 415
QY 397 KPRGEGLDYGVAREWLYLICHEMLNPPYGLFOYSTDNIMLOINPDSSINPDHLSYFH 456
D 416 IPFGEGLDYGVAREWFLLSHEVLNPMYCLFEYAGKDNCLQINPASYINPDHLKYFR 475
QY 457 FYVRIMGLAVFHGHYINGFTVPFYKQLLGPQISLDESVDPELHKSILVLENDITPV 516
D 476 FGRIEALFAGKIDTGFSLPFYKRIILKPVGLKDLSDIDPEFYNLSIWKNNISEC 535
QY 517 -LDHPTCEVHNAGRILOHELKPNGRNVPVTENKKEVRLVYVWRFMRGIEAQLQK 575
D 536 DLEMVPSVDKEILGKIDKPNNGNLTVEENKEEYIRMAEWRSLRSGVEEQTAFFE 595
QY 576 GFNELIPOHLKPFDPQKELELIIGLDKIDLDNWKSNTRLKHCVADSNIVRWFQOAVTF 635
D 596 GFNEILPOQYLOFYDAKELEVLCCQOQIDLDNWRHAIYRHARTSKQIMWFQVKEI 655
QY 636 DEERRARLLQFVTGSTRVPLQFKALQSGTGAAGPRLFTIHLIDANTNLKATCFNRI 695
D 656 DNEKRMRLQFVTGTCTRLPVGGFADLMGNSN---GPKFCIEKV-GKENWLPKRSHTCFNRL 711
QY 696 DIPPYESYEKLYEKLITAVEETCGFAVE 723
D 712 DLPFYKSYEQLEKELFAIBETEGFQOE 739

RESULT 12
US-10-097-534-10
; Sequence 10, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; TITLE OF INVENTION: MATURATION
; FILE REFERENCE: PLV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/308,958
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 739

Query Match 32.1%; Score 1245.5; DB 12; Length 739;
Best Local Similarity 39.8%; Pred. No. 9,8e-105;
Matches 274; Conservative 99; Mismatches 180; Indels 135; Gaps 18;
US-10-182-936A-89
```

```
QY 219 RPHGOSPE-----LPEGYEORTTVQGVYFLHTGTGVTWHDPRIP-----RDLSNV 266
D 109 SFASATSEDSGSGSLPPTNTNTNTSEGTSLIPL-----TISGGSGP-RPLNPV 159
QY 168 MEPEPAPYTDSTGAAAGGNCRFVESPSQDQLQOARLNPDPVRSGL-----QTPQNR 219
D 160 TQAPLP-----PGWEQVR-----DQGRVYVVDHVEXRTTMDR 192
QY 220 RHGHOSPELPEGEORTTVQGVYFLHTGTGVTWHDPRIP----- 260
D 193 PE-----PLPGWEARVDNMGRIYVDFHTTTRTTWQRTLESVRNVEQWQORQLOQAM 247
QY 261 -----RDL-----NSVNCDELPLPGWEVRSTVGRIYVVDHNNRTTQTDPR 305
D 248 QCFNQRFYGNQDLFATQSKEFDPLGPLPGWEKRTDSNGRVFVNHNTRITQWEDPR- 306
QY 306 HHMMHQCOLKEPSQPLP-----PSGSLDEBELPAQR 339
D 307 -----SQGLNE--KPLPEGWEMRTVGPYFVDHNNRTTYIDRTGKALDNGPOIA 359
QY 340 YERDLVQKLKVLK-----HELSQLQOPQAGHCRIEVSREEIFESYQIMKVRKDKRLMV 396
D 360 YVRDFKAKVQYFRFCQQLAMPQ-----HIKITVTRKTLTFEDSFQOIMGFSQDLRRRLMV 415
QY 397 KPRGEGLDYGVAREWLYLICHEMLNPPYGLFOYSTDNIMLOINPDSSINPDHLSYFH 456
D 416 IPFGEGLDYGVAREWFLLSHEVLNPMYCLFEYAGKDNCLQINPASYINPDHLKYFR 475
QY 457 FYVRIMGLAVFHGHYINGFTVPFYKQLLGPQISLDESVDPELHKSILVLENDITPV 516
D 476 FGRIEALFAGKIDTGFSLPFYKRIILKPVGLKDLSDIDPEFYNLSIWKNNISEC 535
QY 517 -LDHPTCEVHNAGRILOHELKPNGRNVPVTENKKEVRLVYVWRFMRGIEAQLQK 575
D 536 DLEMVPSVDKEILGKIDKPNNGNLTVEENKEEYIRMAEWRSLRSGVEEQTAFFE 595
QY 576 GFNELIPOHLKPFDPQKELELIIGLDKIDLDNWKSNTRLKHCVADSNIVRWFQOAVTF 635
D 596 GFNEILPOQYLOFYDAKELEVLCCQOQIDLDNWRHAIYRHARTSKQIMWFQVKEI 655
QY 636 DEERRARLLQFVTGSTRVPLQFKALQSGTGAAGPRLFTIHLIDANTNLKATCFNRI 695
D 656 DNEKRMRLQFVTGTCTRLPVGGFADLMGNSN---GPKFCIEKV-GKENWLPKRSHTCFNRL 711
QY 696 DIPPYESYEKLYEKLITAVEETCGFAVE 723
D 712 DLPFYKSYEQLEKELFAIBETEGFQOE 739

RESULT 12
US-10-097-534-10
; Sequence 10, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; TITLE OF INVENTION: MATURATION
; FILE REFERENCE: PLV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/308,958
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/340,170
; PRIOR FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 739
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-534-10

Query Match      32.1%; Score 1245.5; DB 14; Length 739;
Best Local Similarity 39.8%; Pred. No. 9.8e-105;
Matches 274; Conservative 99; Mismatches 180; Indels 135; Gaps 18;

QY 114 NPSDTDAVGOIVVSL-----QTRDRIG-TGGSVDCRGLLENEGTVYEDSGRPLSCF 167
Db 109 SPATSESDGSGTSLPPNTNTNTSEGATGLIPL-----TISGSGP-RPLNPV 159
QY 168 MEPPAPYTDSTGAAGGCRFVESQDQRLQARLNPDPVRSGL-----QTPQNR 219
Db 160 TQAPLP-----PWEGRV-----DQGRVYVVDHVEXRTTWR 192
QY 220 PHGQSPELPEGYEORTTVQGVYFLHTQGVSTWHDPRIP----- 260
Db 193 PE-----PLPPGWRVNDNMGRYYVDHFRITTTTQWPTLESVRNYEQMLQORSOLOQAM 247
QY 261 -----RDL-----NSVNCDELGPLPPGWEVRSTVSGRIYFVDHNNRTTQTDPR 305
Db 248 QQFNORFIYGNQDLFATSQSKFDPGLPPLPGWEKRTDSNGRVYFVNNTRITOWEDPR- 306
QY 306 HHIMNHQCOLKEPSQPLPL-----PSEGLDEBELPAQR 339
Db 307 -----SQQLNE--KPLPEGWEMRFTVDGIPYFVDHNNRTTIDPRTGKSALDNGPQIA 359
QY 340 YERDLVQKLKVLK---HELSLQOPQAGHCRIEVSREIEESYRQIMKVRPDXLKKRLMV 396
Db 360 YVDFKAKVQYFFWCQOLAMPQ---HIKITVTRKTLFEDSFQOIMSFSDQLRRRLV 415
QY 397 KFRGEGLDYGGVAREWMLYLLCHEMLNPYGLFOYSTDNIMYLQINPDSSINPDHLSYFH 456
Db 416 IPPGEGLDYGGVAREWMLYLLCHEMLNPYGLFOYSTDNIMYLQINPDSSINPDHLSYFH 475
QY 457 FVGRIMGLAVFHGHIYNGFTVPFYKQLGKPIQLSDLESVDPELHSLVWILENDITPV 516
Db 476 FIGRFIAMLFGKFDITGSLPFYKRIKLNKPVGLKXLESIDPEFYNSLIWVKNENIEC 535
QY 517 -LDHTFCVHNFAFRILQHELKPNGRNVPVTENKKEYVRLYVNWFRMGEIAQFLALQK 575
Db 536 DLEMYSVDKEILGEIKSHDLKPNGGNLYTEENKEEYIRMAEWELSRGVEEQTAFFE 595
QY 576 GFNELIPQHLKPPDQKELELIIGGLDKIDLNDWKSNTRLKHCVADSNIVRWFWQAVETF 635
Db 596 GFNEILPQQYLOFYDAKELEVLCCGMEIDLNDWQRAHYRHYARTSKQIMWFQVKEI 655
QY 636 DEERRARLLQFVTGSTRVPLQGFALQSGTGAAGPRLFTIHLIDANTDNLKHAHTCFNRI 695
Db 656 DNEKRMELLOFVTGTCRLPVGGFADLMGNS---GPOKFCIEKV-GKENWLPKSHTCFNRL 711
QY 696 DIPPEYSEKLYEKLTAVEETCGFAVE 723
Db 712 DLPPYKSYEQLEKLLFAIETEFGQGE 739

RESULT 14
US-10-374-979-89
; Sequence 89, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
```



```
Db      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
573 GRVAGLAVFHGKLLDGGFFIRPFYKMLGKQITLNDMESVDSEYNSLKWILENDPTE-LD 631
QY      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
519 HTFCVEHNAFGRILQHELKPNGRNVPVTBENKKEYVRLVYVNRWFMGIEAQFLAQKGN 578
Db      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
632 LMFIDEENFGQTYQVDLKNPNGSEIWTBENKKEYIDLVIQWRFNVRVQKQWNAFLEGFT 691
QY      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
579 ELIPQHLKPFDOKELELIIGGLKIDLNDWKSNTRLKHCVADSN-IVRWFQAVETFE 637
Db      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
692 ELLPIDLIKIFDENELELLMCGLDVDVNDWRQHSYKNGYCNHPVIOFWKAVLLMDA 751
QY      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
638 EREARLLOFVTGSTRYPLOGFKALQSTGAAGPRLETHLIDANTDNLKHAHTCFNRDI 697
Db      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
752 EKIRLLOFVTGSTRVPMNGFAELYSN---GPQLFTISQW-GSPEKLPRAHTCFNRDL 807
QY      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
698 PPYESYKLYEKLLTAVEETCGF 720
Db      ||: ||||| :| | ||| :| | ||| :| | ||| :| | ||| :| | ||| ||
808 PPYETEDLREKLLMAVENAQGF 830
```

Search completed: September 21, 2004, 08:04:23  
Job time : 99.4109 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:39:16 ; Search time 30.0014 Seconds  
(without alignments)  
2398.267 Million cell updates/sec

Title: US-10-009-945-4

Perfect score: 4038

Sequence: 1 MSNPGRRNGPVKLELTWLC.....EKLYEKLTAETFCGPAVE 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:\*\*

1: pir1:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1585.5	39.3	767	2 T37545	ubiquitin-protein
2	1564	38.7	766	1 S65562	ubiquitin-protein
3	1521	37.7	815	2 T49744	probable ubiquitin
4	1484	36.8	809	1 S43217	ubiquitin-protein
5	1418	35.1	786	2 T39585	ubiquitin-protein
6	1364.5	33.8	820	2 T46412	ubiquitin-protein
7	1355.5	33.6	887	2 S70642	ubiquitin ligase N
8	1177	29.1	708	2 T83136	NBDD-4 ORF - mouse
9	1046.5	25.9	671	2 T37900	probable ubiquitin
10	858	21.2	3227	2 T37954	probable ubiquitin
11	846.5	21.0	3839	2 T49759	related to TOM1 pr
12	805.5	19.9	1126	2 T01431	ubiquitin-protein
13	803	19.9	4056	2 H96599	ubiquitin-protein
14	789	19.5	3268	2 S69623	protein F14J16.10
15	709.5	17.6	889	2 T20274	hypothetical prote
16	626	15.5	310	2 I52646	hypothetical prote
17	596	14.8	221	2 T51896	DNA binding protei
18	534.5	13.2	874	2 A38920	hypothetical prote
19	465	11.5	807	2 T40821	86-associated prot
20	460.5	11.4	892	2 S57055	probable ubiquitin
21	454	11.2	873	2 H85134	probable membrane
22	449	11.1	1142	2 T46155	polyubiquitin-like
23	445.5	11.0	1054	2 B38919	hypothetical prote
24	444.5	11.0	1083	2 A38919	hypothetical prote
25	418.5	10.4	1029	2 T38951	hypothetical prote
26	407.5	10.1	4861	2 S71752	probable ubiquitin
27	403.5	10.0	1066	2 T26753	giant protein p619
28	387	9.6	1001	2 T20373	hypothetical prote
29	371	9.2	4836	2 T14346	herc2 protein - mo

30 360 8.9 655 2 T41750 probable ubiquitin  
31 358 8.9 757 2 T05688 hypothetical prote  
32 355.5 8.8 1502 2 T48309 hypothetical prote  
33 342 8.5 910 2 S64155 hypothetical prote  
34 297.5 7.4 2899 2 T21546 hypothetical prote  
35 297.5 7.4 2915 2 G87867 protein P36A2.13 [  
36 286.5 7.1 959 2 T14761 hypothetical prote  
37 281 7.0 889 2 S22659 hypothetical prote  
38 278.5 6.9 1483 2 S30015 hypothetical prote  
39 262.5 6.5 632 2 T38617 hypothetical prote  
40 245.5 6.1 2761 2 T29285 probable ubiquitin  
41 224.5 5.6 2895 2 T08437 hypothetical prote  
42 220.5 5.5 1630 2 T00390 hyperplastic discs  
43 207 5.1 472 2 B56954 KIAA0614 protein -  
44 204 5.1 1256 2 JE0209 yes-associated pro  
45 197 4.9 1171 2 T42372 brain-specific ang  
probable guanylate

#### ALIGNMENTS

##### RESULT 1

T37545  
ubiquitin-protein ligase (EC 6.3.2.19) pub1 [validated] - fission yeast (Schizosaccharomy  
N; Alternate names: ubiquitin ligase Pub1  
C; Species: Schizosaccharomyces pombe  
C; Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Jun-2002  
C; Accession: T37545; T48655  
R; McLean, J.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, September 1997  
A; Reference number: Z21722  
A; Accession: T37545  
A; Status: preliminary; translated from GB/EMBL/DBJ  
A; Molecule type: DNA  
A; Residues: 1-767 <NCL>  
A; Cross-references: EMBL:Z99161; PIDN:CAB16207.1; GSPDB:GN00066; SPDB:SPAC11G7.02  
A; Experimental source: strain 972h-; cosmid c11G7  
R; Saleki, R.; Jia, Z.; Karagiannis, J.; Young, P.G.  
Mol. Gen. Genet. 254, 520-528, 1997  
A; Title: Tolerance of low pH in Schizosaccharomyces pombe requires a functioning pub1 ubi  
A; Reference number: Z07985; MUID:97340937; PMID:9197411  
A; Accession: T48655  
A; Status: preliminary; translated from GB/EMBL/DBJ  
A; Molecule type: DNA  
A; Residues: 1-767 <SAL>  
A; Cross-references: EMBL:U62795; NID:G2262192; PIDN:AAB63350.1; PID:G2262193  
A; Experimental source: strain J227  
C; Genetics:  
A; Gene: pub1; SPDB:SPAC11G7.02  
A; Map position: 1  
A; Introns: 6/2; 14/1; 62/2  
C; Function: <CYC>  
A; Description: EC 6.3.2.19 [validated, MUID:96205868]  
C; Function: <CYC>  
A; Description: Involved in of the mitotic activating tyrosine phosphatase cdc25 [validate  
C; Function: <TOU>  
A; Description: required for low pH-tolerance [validated, MUID:97340937]  
C; Superfamily: yeast ubiquitin-protein ligase, WW repeat homology  
C; Keywords: cell cycle control; ligase  
F; 205-242/Domain: WW repeat homology <WW1>  
F; 288-325/Domain: WW repeat homology <WW2>  
F; 345-382/Domain: WW repeat homology <WW3>

Query Match 39.3%; Score 1585.5; DB 2; Length 767;

Best Local Similarity 41.3%; Pred. No. 9.8e-107;

Matches 340; Conservative 121; Mismatches 231; Indels 131; Gaps 17;

QY 1 MSNPGRRNGPVKLELTWLC...EKLYEKLTAETFCGPAVE 748

Db 1 MSNPGRRNGPVKLELTWLC...EKLYEKLTAETFCGPAVE 748

QY 61 WQVHVDLYIGKSDSVTISVWVNHKKHKKQAGFLGCVFLLSNRLKXDTGYQL--DLG 118

Db 56 WNETFEVNTDNTSTIAIQVFDQKPF-KKKGQGLGVINLRVGVLDLAIGDEMTRDLK 114  
 QY 119 KLGPNNDTVRGQIVVSIQ----- 137  
 Db 115 K--SNENTVVGKIIINLSTTAOSTLOVPSSAASGARTQRTSITNDPQSSQSSVSSENPA 172  
 QY 138 -----SRDRIGTGGQVDDCSRLFDN-----DLPDGWEERTASGRIOYLNIHITRTTQ 184  
 Db 173 SSRAGSPTRDNPAPASPASSEPTFSFEDQYGLRPFGEWERTDNLGRTYVVDHNTSTT 232  
 QY 185 WERTPT-----RPASEYSSGRPLSCFVDENTPISGTNGATCGOSSDPLAERRVRSORH 238  
 Db 233 WIRPNLSSVAGAAAELHSSAS--SANVTGQVQPSSSNAA-----RTEASVLT--- 279  
 QY 239 RNYMSRTHLTPDLPPEGYEORTTQCGQVYFLHTQTVSTWHDPRVPRDLSNIN----- 292  
 Db 280 ----SNATTAGSGELPPGWEQRYTPEGKPYFVDHNTTRITTTWDPRRQYIRSYGGPNNAT 335  
 QY 293 -----CEELGPLPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQNLKQDQ 347  
 Db 336 IQQQPVSQLGPLSGWEMRLTNTARVYFVDHNTKTITWDDPRLPSSL-----DQ 384  
 QY 348 QQQQVVSLLCPDDTECLTVPRYKRDVLQKILRQELSQ--QPQAGHCRIEVSREIPEE 405  
 Db 385 -----NVQYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHIIFED 425  
 QY 406 SYRQVWVKORPKDLWKRLMKIFRGEGLDYGGAAREWLYLLSHEMLNPPYGLFOYSRDDIY 465  
 Db 426 SYAELMRQASATDLKRLMTKFDGEDGLDYGGLSREYFLLSHEMFPNPFYCLPEYSSVDNY 485  
 QY 466 DPDLHNSLVWILENDITGVLDHTFCVEHNAYGEIIOHELKPNKGIPIVNEENKEYVRLY 585  
 Db 546 DAEYRSVLWILDNDITGVLDHTFVEDNCFGEVVTIDLPKNGRNIETVEENKREYVDLV 605  
 QY 586 VNRFLRGIEAQLKQGFNEVPIQHLKLTDFEKELELIIICGLGKIDVNDKVNTRLKH 645  
 Db 606 TVW-IQRKEIEQNAPHEGFSELIPELINVDERELELLIGGISEIDMEDWKKHVDYS 664  
 QY 646 CTSDNSITVKWFWKAVFFDEERRARLQVGTSSRVPIQGFALOGAAGPRFTTHQIDA 705  
 Db 666 YSENDQIIKWFWELMDWESNEKSRLLQTTGTSIPVNGFVXDLQSGDGPRTTEKAGE 724  
 QY 706 CTNNLPKAHTCFNRIDIPPYESYEKLYEKLTAIBETCGFAVE 748  
 Db 726 -PNKLPAHTCFNRDLDPPTYSKDLHLKLSIAVEETIGFGQE 767  
 RESULT 2  
 S66562  
 Ubiquitin-protein ligase (EC 6.3.2.19) - fission yeast (Schizosaccharomyces pombe)  
 N:Alternate names: E6-AP-like protein ubiquitin ligase  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002  
 C:Accession: S66562; T45159  
 R:Nefsky, B.; Beach, D.  
 EMBO J. 15, 1301-1312, 1996  
 A>Title: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of cdc25  
 A:Reference number: S66562; MUID:96205866; PMID:8635463  
 A:Accession: S66562  
 A>Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-766 <NEP>  
 R:Nefsky, B.S.; Beach, D.  
 submitted to the EMBL Data Library, August 1996  
 A:Description: Publ acts as an E6-AP-like protein ubiquitin ligase in the degradation of  
 A:Reference number: Z22935  
 A:Accession: T45159  
 A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
 A:Residues: 1-766 <NE2>  
 A:Cross-references: EMBL:Y07592; PIDN:CAA68867.1  
 C:Genetics:  
 A:Gene: publ  
 C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
 C:Keywords: ligase  
 F:205-242/Domain: WW repeat homology <WW1>  
 F:288-325/Domain: WW repeat homology <WW2>  
 F:345-382/Domain: WW repeat homology <WW3>  
 Query Match 38.7%; Score 1564; DB 1; Length 766;  
 Best Local Similarity 41.1%; Pred. No. 3 Se-105;  
 Matches 338; Conservative 121; Mismatches 232; Indels 132; Gaps 18;  
 QY 1 MSNPRRNRGPKRLRTVLCAKVLKDKDFRLLPDPFPAKVVDGSGOCHSTDTVKTLDPK 60  
 Db 1 MSNSAQSR---RIRVTIIVAADGLYKRDVFRFPDPFPAVLTVDGE-QTHTTTAKKTLNPY 55  
 QY 61 WNGHYDLYIGKSDSVTISVWVNHKKIHKQAGFLGCVRLLSNAINRLKDTGYORL--DLC 118  
 Db 56 WNETFEVNTDNTSTIAIQVFDQKPF-KKKGQGLGVINLRVGVLDLAIGDEMTRDLK 114  
 QY 119 KLGPNNDTVRGQIVVSIQ----- 137  
 Db 115 K--SNENTVVGKIIINLSTTAOSTLOVPSSAASGARTQRTSITNDPQSSQSSVSSENPA 172  
 QY 138 -----SRDRIGTGGQVDDCSRLFDN-----DLPDGWEERTASGRIOYLNIHITRTTQ 184  
 Db 173 SSRAGSPTRDNPAPASPASSEPTFSFEDQYGLRPFGEWERTDNLGRTYVVDHNTSTT 232  
 QY 185 WERTPT-----RPASEYSSGRPLSCFVDENTPISGTNGATCGOSSDPLAERRVRSORH 238  
 Db 233 WIRPNLSSVAGAAAELHSSAS--SANVTGQVQPSSSNAA-----RTEASVLT--- 279  
 QY 239 RNYMSRTHLTPDLPPEGYEORTTQCGQVYFLHTQTVSTWHDPRVPRDLSNIN----- 292  
 Db 280 ----SNATTAGSGELPPGWEQRYTPEGKPYFVDHNTTRITTTWDPRRQYIRSYGGPNNAT 335  
 QY 293 -----CEELGPLPPGWEIRNTATGRVYFVDHNNRTTQTPRLSANLHLVLRNQNLKQDQ 347  
 Db 336 IQQQPVSQLGPLSGWEMRLTNTARVYFVDHNTKTITWDDPRLPSSL-----DQ 384  
 QY 348 QQQQVVSLLCPDDTECLTVPRYKRDVLQKILRQELSQ--QPQAGHCRIEVSREIPEE 405  
 Db 385 -----NVQYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHIIFED 425  
 QY 406 SYRQVWVKORPKDLWKRLMKIFRGEGLDYGGAAREWLYLLSHEMLNPPYGLFOYSRDDIY 465  
 Db 426 SYAELMRQASATDLKRLMTKFDGEDGLDYGGLSREYFLLSHEMFPNPFYCLPEYSSVDNY 485  
 QY 466 DPDLHNSLVWILENDITGVLDHTFCVEHNAYGEIIOHELKPNKGIPIVNEENKEYVRLY 585  
 Db 546 DAEYRSVLWILDNDITGVLDHTFVEDNCFGEVVTIDLPKNGRNIETVEENKREYVDLV 605  
 QY 586 VNRFLRGIEAQLKQGFNEVPIQHLKLTDFEKELELIIICGLGKIDVNDKVNTRLKH 645  
 Db 606 TVW-IQRKEIEQNAPHEGFSELIPELINVDERELELLIGGISEIDMEDWKKHVDYS 664  
 QY 646 CTSDNSITVKWFWKAVFFDEERRARLQVGTSSRVPIQGFALOGAAGPRFTTHQIDA 705  
 Db 666 YSENDQIIKWFWELMDWESNEKSRLLQTTGTSIPVNGFVXDLQSGDGPRTTEKAGE 724  
 QY 706 CTNNLPKAHTCFNRIDIPPYESYEKLYEKLTAIBETCGFAVE 748  
 Db 726 -PNKLPAHTCFNRDLDPPTYSKDLHLKLSIAVEETIGFGQE 766  
 RESULT 3

T49744  
probable ubiquitin-protein ligase [imported] - Neurospora crassa  
N/Alternate names: protein B24B19.160  
C/Species: Neurospora crassa  
C/Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 18-Aug-2000  
C/Accession: T49744  
R/Schults, U.; Aign, V.; Hohns, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A/Reference number: Z25022  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-815 <SCH>  
A/Cross-references: EMBL:AL356102; GSPDB:GN00116; NCSP:B24B19.160  
A/Experimental source: BAC clone B24B19; strain OR74A  
C/Genetics:  
A/Gene: NCSP:B24B19.160  
A/Map position: 6  
A/Introns: 11/1; 24/1; 59/2; 110/1; 783/2  
C/Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
F/239-276/Domain: WW repeat homology <WWR1>  
F/334-371/Domain: WW repeat homology <WWR2>  
F/393-430/Domain: WW repeat homology <WWR3>  
Query Match 37.7%; Score 1521; DB 2; Length 815;  
Best Local Similarity 38.8%; Pred. No. 5.1e-102; Indels 170; Gaps 20;  
Matches 332; Conservative 124; Mismatches 229;  
16 LTVLCANLVKDPFRLPDPFPAKVVDGSGCHSTDTVKNTLDPKNOHYDL-----I 69  
9 LPVIAADGLYKRDVPRFPDPAVATNGE-QTKTQVSKRTLNPNVNEHDFWLIQSKV 67  
70 GKSDSVTVSVNHHKHKHKGAGFLGCVRL-LSNAINELKDTGYQLR---DLCKLGPNDND 126  
68 NEDSLAVQVDFQKAF-KKKQGGFLGVINRGVDELAPDAEDQMLTRDLKK---STDNL 124  
127 TVRGQIVVSLQ-----SRDRI-----GTG 145  
125 VVHGKLIINLTNTATWSRLGPPSSSRPSLLTPQSSVINDRANERPSSAMSGPNTA 184  
146 GOVVDCSR-----LPDND---LPDG 162  
185 NMTLASSPASLAVSSSSTAPTGTNGTAPNTPTLVPAQARHSTLSPEDSMGRLPAG 244  
163 WEERTASGRQYLNHTRTQWERPTPASEYSPGRPSCFVDENPTI----- 212  
245 WERREDHLGRYYVDHSRTTSWRPT-----GTGAENRTAEANTQVERQHQRNTL 297  
213 ----SGTNGATCGSSDPRLAERRVRSORHRYNYSRTHLHTP--PDLPGEYQRTTQQG 266  
298 PEDRTGAN-----SPTLQQQAAATANAATMMHTGATTPGTGELPAGWEQRTPEGR 349  
267 VYFLHTQGVSTWHDPRVPRDL-----SNIN-----CDELGLPPEGHEINTATGRV 313  
350 PYFDVHNTRTTTWDPRRQYIRMYGQNNNGTIIQQPVSQGLPPLSGMELTNTARV 409  
314 YFVDHNRNTTQFTDPRLSANLHLVLNRQNLKQOQQQVWSLCPDDTECLTVPRYKRDV 373  
410 YFVDHNTKTTTWDPRLPSSL-----DQ-----NVQYKEDR 442  
374 QKLKILRQLSQOQPOACHCEIEVSREIPESYRQVWKPDPDLWRLMIKFRGEGSLD 433  
443 RKLIFRSQ-PAMRMSQCCHIKYRRSHIPEDSFABISROSATDLKRLMIKFDGEDGLD 501  
434 YGVVAREWLYLLSHEMLNPPYGLFYQYGRDIIYTIQINPDGAVNPEHLSYFFHVGIRMGMA 493  
502 YGGLSREFFFLSHMEMNPFYCLEYSADHNTYTIQINPHSGINPEHLNYPKFTGRVVGLA 561  
494 VFHGHYDGGTLPFYKQLGKSTLDDMELVDPDLNSLWILENDITGVLDHTFCVEH 563  
562 IFRFRFLDAPFIALGYKWLKAVSLADMEGVADPFRSLQWMLDNDITDVLDAFTFSD 621  
554 NAYGEIIQHELKPKNGKSPVNVNEENKVEYVRLVYVNWRLRGIEAQFLALQKGFNEVFPQHL 613

## RESULT 4

S43217  
ubiquitin-protein ligase (EC 6.3.2.19) RSP5 - yeast (Saccharomyces cerevisiae)  
N/Alternate names: E6-AP-like protein ubiquitin ligase; protein YER125w; PUB1 protein  
C/Species: Saccharomyces cerevisiae  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002  
C/Accession: S43217; S50628; S70050  
R/Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor, I.  
submitted to the EMBL Data Library, February 1993  
A/Reference number: S30812  
A/Accession: S43217  
A/Molecule type: DNA  
A/Residues: 1-809 <MUL>  
A/Cross-references: GB:U18916; EMBL:L111119; NID:G1384128; PIDN:AAC03223.1; PID:G603364  
R/Dietrich, F.S.  
submitted to the EMBL Data Library, December 1994  
A/Description: The sequence of S. cerevisiae cosmid 9781, 8198, 9115, 9981, and lambda c  
A/Reference number: S50628  
A/Accession: S50628  
A/Molecule type: DNA  
A/Residues: 1-809 <DIE>  
A/Cross-references: EMBL:U18916; NID:G1384128; PIDN:AAC03223.1; PID:G603364; GSPDB:GN0000  
R/Hain, C.; Springael, J.Y.; Volland, C.; Haguenauer-Tsapis, R.; Andre, B.  
Mol. Microbiol. 18, 77-87, 1995  
A/Title: NP11, an essential yeast gene involved in induced degradation of Gap1 and Fur4  
A/Reference number: S70050; MUID:96154942; PMID:8596462  
A/Accession: S70050  
A/Status: nucleic acid sequence not shown; not compared with conceptual translation  
A/Molecule type: DNA  
A/Residues: 1-101 <HEI>  
C/Genetics:  
A/Gene: SGD:RSP5; PUB1; NP11; MIPS:YER125W  
A/Cross-references: MIPS:YER125w; SGD:S0000927  
A/Map position: 5R  
C/Function:  
A/Description: involved in endocytosis of GAP1 protein and FUR4 protein; binds and ubiqui  
C/Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
C/Keywords: ligase  
F/229-266/Domain: WW repeat homology <WML>  
F/331-368/Domain: WW repeat homology <WM2>  
F/387-424/Domain: WW repeat homology <WM3>  
Query Match 36.8%; Score 1484; DB 1; Length 809;  
Best Local Similarity 37.1%; Pred. No. 2.4e-99;  
Matches 324; Conservative 125; Mismatches 224; Indels 200; Gaps 20;  
11 PVKRLRVLCAKLVKDPFRLPDPFPAKVVDGSGCHSTDTVKNTLDPKNOHYDL-YI 69  
2 PSSISVKLVAAESLYKRDVFRSPDPFVAVLTIDGY-QTKSTSAKTLNPNYNETKFDI 60  
70 GKSDSVTVSVNHHKHKHKGAGFLGCVRL-LSNAINRLKND-----TGVR-----LDLC 118  
61 NENSILTIQVFOQKXF-KKKDQOQFLGVNVRVGVLDHLEDTATSSGRPRETTIRDLK 119  
119 KLGPNDNDTVRGOIVV-----SLQSRDRIGTGQVVDCSR--- 153  
120 K-SNDGMVSGRLIVLVLSKLPSSSPHQAQPSGHTASSSTTSSTRTRNGHSTSTRNHS 177

154 QY -----LFDND---LPDGHNE 165  
 178 DB TGHPSRGTAQAVESTLQSCGTAATNTATTSHRSTNSTSATQYSSFDQYGRLLPPGWER 237  
 166 QY RTASGRICVLAHITTTTOWERPTPASE-----YSSPGRPLSCFVDENT 210  
 238 DB RTDNFORTYVDHNTTITTKRPTLDQTEAERGNQLNANTELERRQHRGRTLPGSSSDNS 297  
 211 QY -----PISG-----TNGATCGOSSDPRLAERRVRSQRHNNYMSHTLHT 249  
 298 DB SVTVQVGGGNSIPVNGAAAAAFAATGTTSGLG----- 331  
 250 QY PPDLPGEYQRTTQOQVYFLHTQGVSTWHDPRVPRDLSNIN-----CEELGFL 299  
 332 DB --ELPSGWEQRFPPEGRAYFVDHNTTITTVWPRQOYRTTYGPTNTTIOQPVSLGFL 389  
 300 QY PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNLKQDQOQOQVSLCPDD 359  
 390 DB PSCEMELNTARVYFVDHNTKTTTWDPRLPSSL-----DQ----- 426  
 360 QY TECLTVPRYKRDVLQKILRQELSOQOQ-----AGHCIEVSRBEIPEESYRQVWKMXP 415  
 427 DB ---NVPQYKRDPRKVIYFR-----SQPALRILPGQCCHKVRRKNIFEDAIQEMRQTP 477  
 416 QY KDLWKRLMIKFRGEGLDYGVAREMLYLLSHEMLNPPYGLFYQSRDDIYTLQINPDSAV 475  
 478 DB EDLKKRLMIKFDGEGLDYGVSRBFFLLSHEMFNFYCLFEYSAYDNYTIQINPNSGI 537  
 476 QY NPEHLSYFHVGRIMGVAFHGYIYDGGFTLPFYKOLLGKSTLDDMELVDPDLHNSLVW 535  
 538 DB NPEHLNYFKYIGRVLGVGFHRRFLDAFFVGLYKMLRKKVYLQDMGVDAEYVNSLW 597  
 536 QY ILNDITGVLDHFTCVENHAYGELIQHELKPNKSIPIVNEENKKEYVRLVYVNWRLRGIE 595  
 598 DB MLENSIDGVLDLIFSADDERFGEVTVVDLKPDRNIEVTDGNKKEVVELYTKQWIDRVQ 657  
 596 QY AQELALQKGFNEVPIQHLKTFDEKELELIICGLKIDVNDWKVNRHLKCTPDNSIVKW 655  
 658 DB EQKAFMDGFNELIPEDLVTVDFERELELLIGIAIEDIEDMKKTDYRGYQESDEVIOQ 717  
 656 QY FWKAVEFDEBRARILQFTVGSRRVPLQGFALQAGAPRLFTTHQIDACTNNLPKXHT 715  
 718 DB FWKCVSEWDEQARLLQFTTGSRIPVNGFDLQSGDGRPRFTIEKAGE-VQQLPKSHT 776  
 716 QY CFNRIDIPPPYESYEKYLEKLLTAIBETCGFAVE 748  
 777 DB CFNRVDLPQYVDYDSMKQKLTLAVEETIGFGQE 809  
 RESULT 5  
 T39585  
 ubiquitin protein ligase - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 18-Aug-2000  
 C:Accession: T39585  
 R:Voickaert, G.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, August 1997  
 A:Reference number: Z21865  
 A:Accession: T39585  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-786 <VOL>  
 A:Cross-references: EMBL:Z99759; PIDN:CABL6903.1; GSPDB:GN00067; SPDB:SPBC16E9.11c  
 A:Experimental source: strain 972h; cosmid c16E9  
 C:Genetics:  
 A:Gene: SPDB:SPBC16E9.11c  
 A:Map position: 2  
 A:Introns: 60/2  
 C:Superfamily: yeast ubiquitin-protein ligase; WW repeat homology  
 F:236-273/Domain: WW repeat homology <WWR1>  
 F:306-343/Domain: WW repeat homology <WWR2>  
 F:364-401/Domain: WW repeat homology <WWR3>

Query Match 35.1%; Score 1418; DB 2; Length 786;  
 Best Local Similarity 36.7%; Pred. No. 1.4e-94;  
 Matches 316; Conservative 127; Mismatches 217; Indels 202; Gaps 20;  
 10 GPVKRLTVLCARNLVKKQFFRLPDPFAKVVVDGSCQCHSTDTVKNTLDPKKNQHYDLVI 69  
 4 GAKRVRFYIIVAADGLSKRDLFRQDPFAITVDGE-QTHTKVVKSVNPVWNEGEFVTV 62  
 70 GKSDSVTISVWVHKIKKKGAGFLGCVRLLSNAINRLKDTGYORLDL-----CKLGP 122  
 63 KPSVISIRLFDOKKP-KKQOQFLGLVSF-----RMREVGSFRNREVSLTRPLKSS 115  
 123 NDNDTVRGQIVV----- 134  
 116 TTNLVLGNLKVAPSKIRAPAGNHSSTTANRTTSTPTTTTARTTRTPRPTATTNS 175  
 135 -----SIQSRDRIGTGG-----QVVDCSRLPDND----- 158  
 176 QSTNSNTRNTSAATNGTGTGAGTGASHRSSPVTRQTNNTSALNSNAHIMSSEFQY 235  
 159 --LPDGWEERTASGRIOYLNHITRTQWERPTRPASEYSSPGRPLSCFVDENTPISGNT 216  
 236 GRLLPPGWERRADSLGRYYVDHNTTITW---TRPAS-----STNPVHNT- 277  
 217 GATCGQSSDPLAERRVRSQRHNNYMSRTHLHTPP-----DLPEGYEORTTQ 263  
 278 -----SSD-----SQR-LNHQNR---HLPDSDNPSLMQSDSGNDLFFGWEMRYTD 318  
 264 QGVYFLHTQTVSTWHDPR-----VPRDLNINCEELGPLPPGWEIR 306  
 319 TGRPYFVDHNTTITWVDPRNPLVRPNNGSGSTVGLMQPQSLSH-----LGPLPSGWEIR 373  
 307 NTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNLKQDQOQOQVSLCPDTECLTYP 356  
 374 LTNSARVYFVDHNTKTTTWDPRLPSSAL-----DQD-----VP 406  
 367 RYKEDLVQKILRQELSOQOQAGHCRIEVSRBEIPEESYRQVWKMXPDKLWRLMIKF 426  
 407 QYKCDFRKRLYFASQ-PGMRPLPGQCNVVRDRHIFEDSYAEMRYSAHDLKRLMIRF 465  
 427 RGEGLDYGVARLWYLLSHEMLNPPYGLFYQSRDDIYTLQINPDSANVPEHLSYFHV 486  
 466 DGEDGLDYGSLSRBFFLLSHKMPDPIYCLFEYSADVNTYTLQINPSSINPEHLYNRFI 525  
 487 GRIMGMAVFGHYIDGGFTLPFYKOLLGKSTLDDMELVDPDLHNSLVWLENDITGVLD 546  
 526 GRVIGLAIFHRFLDAFFVSVLYKLLRKKVSLADMESIDAEFYRSLKWLWLENDITGILD 585  
 547 HTFCVHNAYGEITQHCLKPNKSIPIVNEENKKEYVRLVYVNWRLRGIEAQFLAQKGFN 606  
 586 LTFVSEEDHFGVEVTVELLINGENIEVTEENKKYVDLVTEWVSRKRVQOQFNAYSGFV 645  
 607 EVIPOHLKLTDEKELELIICGLKIDVNDWKVNRHLKCTPDNSIVKWFKVAEFFDEE 666  
 646 ELVSPDLVNVDFERELELLIGISDVEDWKSHTYRTVIATDPVTKRWFEIAGWNE 705  
 667 RRARLLQFTVGSRRVPLQGFALQAGAPRLFTTHQIDACTNNLPKXHTCFNRIDIPPE 726  
 706 DRSLKLQFATGTSRIPVNGFDLQSGDGRPKFTIEKA-GTPDQLPVAHTCFNRLLDLPYP 764  
 727 SYEKLKYLEKLLTAIBETCGFAVE 748  
 765 SKDTLHEKLSLAVENTVGVFGNE 786  
 RESULT 6  
 T46412  
 ubiquitin-protein ligase (EC 6.3.2.19) NEDD4 - human (fragment)  
 N:Alternate names: hypothetical protein DKF434P422.1  
 C:Species: Homo sapiens (man)  
 C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 03-Jun-2002  
 C:Accession: T46412

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Accession: Z23034  
A:Reference number: Z23034  
A:Accession: T45412  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-820 <AAA>  
A:Cross-references: EMBL:AL137469  
A:Experimental source: adult testis; clone DKFZp434P2422  
C:Genetics:  
A:Gene: GDB:NEDD4  
A:Note: DKFZp434P2422.1  
C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin  
C:Keywords: ligase  
F:142-179/Domain: WW repeat homology <WWR1>  
F:342-379/Domain: WW repeat homology <WWR2>  
F:393-430/Domain: WW repeat homology <WWR3>  
F:489-814/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 33.8%; Score 1364.5; DB 2; Length 820;  
Best Local Similarity 37.1%; Pred. No. 1.2e-90;  
Matches 315; Conservative 115; Mismatches 221; Indels 199; Gaps 25;

QY 50 TDTVNTLDPKNQHYDIYIGKSD-SVTISVWNHKKHKKQAGPLGCVR----- 98  
DB 12 TKTIKTLNPKWNEEFYFRVNPNSHRLLFVFDENLRLTRDD---FLGQVDVPLSHLPTED 68  
QY 99 -----LLSNAINRLKDTGYQLDLCKLGPNDNDTVRGQIVVSLOSRRDRTGIG 146  
DB 69 PTMERPYTFKDFLLPRSHKSEVKGLKXAYMPNG-----GQDENSSQORDMEHGW 123  
QY 147 QVDCSRLFDND-----LPDGWEHRTASGRIOYLNHITTTQWERPT----- 189  
DB 124 EYVD-----SNDASQHOEELPPPLPPGQWEEKVDNLGRVYVNHNNRTTQWHRPSLMDV 178  
QY 190 -----RPAEYSSP-----GRPLSCFVDENTPISGTN-- 216  
DB 179 SSESNNTRQINQEAHRRFRSRHISDLPEPSEGQVPEPMET-ISEENVNAGDSL 237  
QY 217 -----GATCGOSSDP-LAE-----RRVR-----SORHRYNMRKTHL----- 247  
DB 238 LALPPPPASPGSRTSQBELSELSRLQITPDSNGEQSSLIQREPSRLSCSVTDAVA 297  
QY 248 --HTPPD-----LPDGWEHRTASGRIOYLNHITTTQWERPT----- 266  
DB 298 EGHLPFGAKSPVRRVAVKDTLSNPQSPQSPYNSPKQHKVQTSFLPPGWMRIAPNGR 357  
QY 267 VYFLHTQTGVSTWHPRPV-----RDLNINCEELGPPGWEIRNTATGRVYFVDHNR 321  
DB 358 PFFIDHNKTTTWEDPRLKPPVHWSKTSLNPNLGLPLPPGWEIRHLDCRTFFIDHNSK 417  
QY 322 TQFTDPLSLANHLVLNRQNLKDDQOQQQVVSCLCPDDETECLTVPR--YKRDVLVOKLIL 379  
DB 418 ITQWEDPRL-----QN-----PAITGPAVPSREFKQKYDYF 449  
QY 380 ROELSQOQPOAGHCHRIEVSREBIEESYQVMKVRPKDLWK-RLMKPRGEGLDYGVA 438  
DB 450 RKLLKPADIPNRFEMKLRHNNIFESYRRIRMSVGRPDVLKARLWIESEKGLDYGVA 509  
QY 439 REWLYLLSHEMLNPYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRMGAVFHG 497  
DB 510 REWFELLSKEMFNPPYGLFEYSATDNYTLQINPNSGLCNEDHLSYFTFGRVAGLAVFHG 569  
QY 498 HYIDGFTLPVYKQLGKSLTLDDELVDPLHNSLWILENDITGLVDHTFCVEHNAYG 557  
DB 570 KLLDGFTRPVPYKMLGKQITLNDNESVDSYNSLRKWLINDPT-ELDLMFCDIENFNG 628  
QY 558 EIIQHELKPGKSIIPVNEENKEYVRLVYNNRFLRGIEAFLAQKGFNEVPOHLLKTF 617  
DB 629 QTYQVDLKPNGSEIMVTNENKREYIDLVQIRFVNRVQKNNAFLGFTTELLPDLIKIF 688  
QY 618 DEKELELLICGLKIDNDWKVNTLKL--HCTPDSNVKVTWKAVEFDEERRARLLQFV 675

DB 689 DENELELLMCGLDGVDVNDHRQHSIYKNGYC-PNHFVIQWFKXAVLMDAEKIRLLQFV 747  
QY 676 TGSRRVPLQGFKALQGAAGPRLLTIHQIDACTNNLPAKHTCFNRRIDIPPYESYEKLYEKL 735  
DB 748 TGTSTRVPMNGFABLYGNGPQLFTIEQWGS-PEKLPRANTCFNRLDLPYVETFDLREKL 806  
QY 736 LTAIETTCGF 745  
DB 807 LMAVENAQGF 816  
RESULT 7  
S70642  
ubiquitin ligase Nedd4 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 18-Aug-2000  
C:Accession: S70642  
R:Staub, O.; Dho, S.; Henry, P.C.; Correa, J.; Ishikawa, T.; McGlade, J.; Rotin, D.  
EMBO J. 15, 2371-2380, 1996  
A:Title: WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na(+) c  
A:Reference number: S70642; MUID:96221297; PMID:865844  
A:Accession: S70642  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-887 <STA>  
A:Cross-references: EMBL:U50842; NID:gl293646; PIDN:AAB48949.1; PID:gl293647  
C:Genetics:  
A:Gene: Nedd4  
C:Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin  
F:54-167/Domain: protein kinase C C2 region homology <KC2>  
F:246-283/Domain: WW repeat homology <WW1>  
F:402-439/Domain: WW repeat homology <WW2>  
F:459-496/Domain: WW repeat homology <WW3>  
F:555-881/Domain: ubiquitin-protein ligase homology <UBI>  
Query Match 33.8%; Score 1355.5; DB 2; Length 887;  
Best Local Similarity 37.1%; Pred. No. 5.8e-90;  
Matches 314; Conservative 131; Mismatches 248; Indels 153; Gaps 25;  
QY 14 LRLTVLCAKMLVKDFRLLPDPFAKV---VDGSOCHSTDTVKNTLDPKNQHYDLY- 68  
DB 77 VRVYVLAGLAKDKILGASDPYRVVTLDPMSGLVTSVQTKIKSLAPKNEEI-LPR 135  
QY 69 -IGKSQSVTISVWNHKKHKKQAGPLGCVR-----LLSNAIN 105  
DB 136 VLPQQRHLLFEVDENLRLTRDD---FLGQVDVPLPPTENPMERPYTFKDFVLHPRSH 192  
QY 106 RLKDTGVQRLDLCKLGPNDNDTVRGQIVVSLOSRRDRTGIGQVVD---CSRL----FDN 157  
DB 193 KSRVKGVLRLKMTYLPNGSDD-----ENAOAELEFGWVVDQPDAAATHLQHPPEPS 246  
QY 158 DLPDGWEHRTASGRIOYLNHITTTQWERPT-----RPAAS 193  
DB 247 PLPPGWEERQDVLGRVYVNHESRTTQWRKPSPEDDLTDENGDIQLQAHGATTTTRQIS 306  
QY 194 E-----YSSPQRPUSCFVDENTPISGTNGATCGSSDP-----RLAE----- 230  
DB 307 EDVGDGPNHESPNWEIREDENTIVSG-QAVQSPSPSGHPDVQVRLAESLTRLTVYGNP 365  
QY 231 ---RRVRSORH---RNYMSRT-----HLHTPPDLPGEVQORTTQOQGVYFLHTQT 274  
DB 366 ATSQPTSSNHSRSGSSQTCIFEEQPTLPVLLPTSSGLPPGWEKQDGRSYYVDHNS 425  
QY 275 GVSTWHPD-----RVP---RDLNINCEELGPPGWEIRNTATGRVYFVDHNNRTT 323  
DB 426 KTTTWSKPTMQDDPRSKIPAHLRGKTFPDSNDLGPLPPGWEERTHTDGRVFFNNHKKT 485  
QY 324 QFTDPLSLANHLVLNRQNLKDDQOQQQVVSCLCPDDETECLTVRYKRDVLVOKLILRQL 383  
DB 486 QWEDPM-----QNVAITGPAP-----YSRDYKRYKVEFRRLK 519  
QY 384 SQOQPOAGHCHRIEVSREBIEESYQVMKVRPKDLWK-FLMTKFRGEGLDYCGVAREWL 442

Db 520 KKQTDIPNKPENKLRANILEDSDYRRIMGYKGRADFLKARLWIEFDGEGKLDYGGVAREWF 579  
 QY 443 YLISHMLNPPYGLFOYS-RDDYITLIQINPDSAV-NPEHLSYFHFVGRIMGMVAVHGHYI 500  
 Db 580 FLISKEMFNPPYGLFEVSATEDNYITLIQINPNSGLCNEDHLSYFKFGRVAGNAVHVKLL 639  
 QY 501 DGGFTLPFFYKQLLGKSIITLDDMELVDPDLHNSLVWILENDITGVLDHTFCVHNAYGEII 560  
 Db 640 DGGFTLPFFYKQLLGKSIITLDDMELVDPDLHNSLVWILENDITGVLDHTFCVHNAYGEII 560  
 QY 561 QHLEKNGKSIIPYNEENKKEYVLYVNWFLRGIEAQLALOKGENEVIPOHLKTFDEK 620  
 Db 699 QHLEKNGKSIIPYNEENKKEYVLYVNWFLRGIEAQLALOKGENEVIPOHLKTFDEK 620  
 QY 621 ELELIICGLGKIDVNDKVNTRLKH-CTPDSNIVKFWKAVFEFFDBERRARLLQFVTGS 679  
 Db 759 ELELLMCGLDVNDVNDWREHTKYNGYSNMQHVIHFWKAVLWMDSEKRIQLLQFVTGS 818  
 QY 680 RVPLOGKALQAGARLFTIHOIDACTNNLPKATCFNRIDIPPYSEYKLYEKLITAI 739  
 Db 819 RVPNGFAELYGSGPQSFVEQW-GTPDKLPRAHTCFNRDLPPYSEFDELWDKLQWAI 877  
 QY 740 EETCGF 745  
 Db 878 ENTQGF 883

RESULT 8  
 NEDD-4 ORF - mouse (fragment)  
 C/Species: Mus musculus (house mouse)  
 C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 18-Aug-2000  
 C/Accession: 183196  
 R/Kumar, S.; Tomooka, Y.; Noda, M.  
 Biochem. Biophys. Res. Commun. 185, 1155-1161, 1992  
 A/Title: Identification of a set of genes with developmentally down-regulated expression  
 A/Reference number: 160167; MUID: 92328780; PMID: 1378265  
 A/Accession: 183196  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-708 <RES>  
 A/Cross-references: GB:D10714; NID:G220508; PID:G220509  
 C/Genetics:  
 A/Gene: NEDD-4  
 C/Superfamily: rat ubiquitin-protein ligase; protein kinase C C2 region homology; ubiquitin  
 F/40-77/Domain: WW repeat homology <WW1>  
 F/196-233/Domain: WW repeat homology <WW2>  
 F/251-288/Domain: WW repeat homology <WW3>  
 F/347-682/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 29.1%; Score 1177; DB 2; Length 708;  
 Best Local Similarity 40.4%; Pred. No. 3.6e-77;  
 Matches 260; Conservative 87; Mismatches 178; Indels 118; Gaps 17;

QY 159 LPDQWEERTASGIQVNIHTRTQWERPT-----RPAE 194  
 Db 42 LPDQWEERTASGIQVNIHTRTQWERPT-----RPAE 194  
 QY 195 -----YSSGRPLSCFVDENTPISG-----TNGATCGQ--S 223  
 Db 102 DVDGPDNRESNWEIVREDENTEYSQAVQSPSGHIDVQTHLASEFNRLAVCGNPAT 161  
 QY 224 SDPRLAERVRVSQRHRYMGRTH-----LHTPPDLPEGYEQRTTQGGVYFL 270  
 Db 162 SQP-----VTSNHSRGGSLQCIPEEQPTLPVLLPTSSGLPPGWEERQDDRGSRYYV 215  
 QY 271 HTQGVSTWDDP-----RVPRDL-SNINCEELGELPGWEIRNTATGRVYFVDHNR 321  
 Db 216 DHNSKTTTWSKPTWQDDPRSKIPAHRLKGTSDNLGELPGWEIRNTATGRVYFVDHNR 275  
 QY 322 TTQFTDPRLSANLHLVLRNQLKXQQQVSLCPDDECTELTPVRYKRLVQKLILRQ 381  
 Db 276 KTQWEDPRL-----QNVAITGP-----AVP-YSRDYKRYKRYEFFR 309

QY 382 ELSDQQQPAQHCHRIEVSREIFEESYRQVMKRPKOLWK-RLMIKFRGEBGLDYGGVARE 440  
 Db 310 KLKQTDIPNKPENKLRANILEDSDYRRIMGYKGRADFLKARLWIEFDGEGKLDYGGVARE 369  
 QY 441 WYLLSHEMLNPPYGLFOYSRDDYITLIQINPDSAV-NPEHLSYFHFVGRIMGMVAVHGHYI 499  
 Db 370 WYLLSHEMLNPPYGLFEVSATEDNYITLIQINPNSGLCNEDHLSYFKFGRVAGNAVHVKLL 429  
 QY 500 IDGGFTLPFFYKQLLGKSIITLDDMELVDPDLHNSLVWILENDITGVLDHTFCVHNAYGEII 559  
 Db 430 LDGGFTLPFFYKQLLGKSIITLDDMELVDPDLHNSLVWILENDITGVLDHTFCVHNAYGEII 559  
 QY 560 QHLEKNGKSIIPYNEENKKEYVLYVNWFLRGIEAQLALOKGENEVIPOHLKTFDE 619  
 Db 489 QHLEKNGKSIIPYNEENKKEYVLYVNWFLRGIEAQLALOKGENEVIPOHLKTFDE 619  
 QY 620 KELELIICGLGKIDVNDKVNTRLKH-CTPDSNIVKFWKAVFEFFDBERRARLLQFVTGS 678  
 Db 549 NELELLMCGLDVNDVNDWREHTKYNGYSNMQHVIHFWKAVLWMDSEKRIQLLQFVTGT 608  
 QY 679 SRVPLQGFKALQAGARLFTIHOIDACTNNLPKATCFNRIDIPPYSEYKLYEKLITAI 739  
 Db 609 SRVPMNGFAELYGSGPQSFVEQW-GTPDKLPRAHTCFNRDLPPYSEFDELWDKLQWAI 877

RESULT 9  
 T37900  
 probable ubiquitin-protein ligase - fission yeast (Schizosaccharomyces pombe)  
 C/Species: Schizosaccharomyces pombe  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 18-Aug-2000  
 C/Accession: T37900  
 R/Rieger, M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, September 1999  
 A/Reference number: Z21752  
 A/Accession: T37900  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-671 <RIB>  
 A/Cross-references: EMBL:AL117390; PIDN:CAB55856.1; GSPDB:GN00066; SPDB:SPAC1805.15c  
 A/Experimental source: strain 972h; cosmid c1805  
 C/Genetics:  
 A/Gene: SPDB:SPAC1805.15c  
 A/Map position: 1  
 A/Introns: 60/2; 105/1; 639/2  
 C/Superfamily: WW repeat homology  
 F/242-279/Domain: WW repeat homology <WWR>

Query Match 25.9%; Score 1046.5; DB 2; Length 671;  
 Best Local Similarity 33.9%; Pred. No. 9.8e-66;  
 Matches 257; Conservative 124; Mismatches 263; Indels 113; Gaps 17;

QY 13 KRLTTLVCAKMLVKKDFRLLPDPFAKVVVDGSGQCHSTDTVNTLDPKWQHLYDIYKGS 72  
 Db 7 EVQLTILHVEGLWKNGLRLSLKPYLLISVD-DDQFKITNVASGTLRLSWGFTKLTVPQ 65  
 QY 73 DSVTISVNNHKKHKKQAGFLGCVRLLSNA---INLKTGYQRLDLCKLGPNDNTR 129  
 Db 66 STLLQLQFDEKQ-KNETSDGFVGLGAAVNSFLPFPNPKDDYKTRITL-----RSPSGSYR 120  
 QY 130 GQIVVSLQSRDRIGTGGQVVD---CSRFLFNDLPDGWEERTASGIQVNIHTRITQW 185  
 Db 121 GS-VVCLFPRSKFPEELPADKSOICTDIIDASGCAWEITDFEGHYVL-----170  
 QY 186 ERPTPEASEYSSPGRPLSCFVDENTPISGNTGATCGGSDPRLAERVRVSQRHRYMGRTH 245  
 Db 171 -----KSP-----QLSVISAISSHEKLENLTPKQKE--VFSQFLFNNSKS 209  
 QY 246 HLHTPPD-----LPEGYEQRTTQGGVYFLHTQTGVSTWHDPRVPRDLNINCEELGPL 299  
 Db 210 SLKINLEYKVIKHLLEHYPLALSVRQV-----AVEXGFL 244  
 QY 300 PEGWEIRNTATGRVYFVDHNRNTTQFTDPR-----LSANLHLVLRNQLKXQQQ 351



submitted to the EMBL Data Library, June 1998  
A:Description: Arabidopsis thaliana chromosome 1 BAC F1707 sequence.  
A:Reference number: Z14334

A:Accession: T01491  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-1126 <YVS>  
A:Cross-references: EMBL:AC003671; NID:g2833627; PID:g3176690; GSPDB:GN00059; ATSP:F1707  
A:Experimental source: cultivar Columbia  
C:Genetics:

A:Gene: ATSP:F1707.15

A:Map position: 1

A:Introns: 118/3; 224/3; 292/1; 373/3; 457/2; 826/3; 875/3; 902/3; 1057/3; 1093/2

C:Superfamily: ubiquitin-protein ligase homology

F:756-1120/Domain: ubiquitin-protein ligase homology <UBI>

Query Match 19.9%; Score 805.5; DB 2; Length 1126;  
Best Local Similarity 41.3%; Pred. No. 6.4e-50;

Matches 177; Conservative 63; Mismatches 142; Indels 47; Gaps

363 LTVPYKRDVLQKILRQSLQQPO--ACHCRIEVSREEIPESYRQVWKMPPKDLWK 420

700 LKAPRL-IDFNKKAYFRSIRHQHDSGLRISVRAYVLEDSYNQLRMSPQDLKG 758

421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480

759 RLNVQFQGEIGIDAGGLTREWYQLLSRVIPDKGALLFTTVGND--TFQPNPNSVYQTEHL 817

481 SYHFVGRINGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLWILEND 540

818 SYKFGVGRMAKALFDQGLLDVYFTRSFYKHLGVKTYTHDIEAVDPDYKYLKWLLEND 877

541 ITGVLDHTFCVE-----HNAY--GEIIOHELKPKNGSIPVNEENKKEYVRLYNNWFLRG 593

878 VSDILDITFSMDADEEKHILYEKTEVTLVLCMCFCLFFWCFFPKCHCIELIILSLMKKYTDY 3854

594 IEAQFLALQGFNEVIPHQLLTKTDFEKELEILICGLKID----- 633

938 IRPQINAFLEGLNELIPRELVSIFNDKELELLISGLPEIDCKUAFYSFLHLFAYSFKV 997

634 -----VNDMKVNTLKHCTPDSNIVKWFVKAFFPDEERRARLLQFVGGSSR 680

998 IITLLSVFFCFLVDLKKANTVTSYTVGSPVIRWVWVKAFFKEDMARFLQFVGTGSK 1057

681 VPLOGFKALQAGAPRLFTTHQIDACTNNLPKHAHTCFNRIDIPPYSYKLYKLLTAIE 740

1058 VPLOGFKALQAGAPRLFTTHQIDACTNNLPKHAHTCFNRIDIPPYSYKLYKLLTAIE 1117

741 ET---CGFA 746

1118 EANEFGGFA 1126

RESULT 13

H96599

Protein F14716.10 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

R:Thetoclogis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: H96599

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-4056 <STO>

A:Cross-references: GB:AE005173; NID:g8778329; PID:AAF79338.1; GSPDB:GN00141

C:Genetics:

A:Gene: F14716.10

A:Map position: 1

Query Match 19.9%; Score 803; DB 2; Length 4056;

Best Local Similarity 40.0%; Pred. No. 6.8e-49;

Matches 177; Conservative 62; Mismatches 143; Indels 60; Gaps

363 LTVPYKRDVLQKILRQSLQQPO--ACHCRIEVSREEIPESYRQVWKMPPKDLWK 420

3617 LKAPRL-IDFNKKAYFRSIRHQHDSGLRISVRAYVLEDSYNQLRMSPQDLKG 3675

421 RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480

3676 RLNVQFQGEIGIDAGGLTREWYQLLSRVIPDKGALLFTTVGND--TFQPNPNSVYQTEHL 3734

481 SYHFVGRINGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLWILEND 540

3735 SYKFGVGRMAKALFDQGLLDVYFTRSFYKHLGVKTYTHDIEAVDPDYKYLKWLLEND 3794

541 ITGVLDHTFCVE-----HNAY--GEIIOHELKPKNGSIPVNEENKKEYVRLYNNWFLRG 593

3795 VSDILDITFSMDADEEKHILYEKTEVTLVLCMCFCLFFWCFFPKCHCIELIILSLMKKYTDY 3854

593 ELKPKNGSIPVNEENKKEYVRLYNNWFLRGIEAQFLALQGFNEVIPHQLLTKTDFEKELEILICGLKID----- 667

3855 ELKPKNGSIPVNEENKKEYVRLYNNWFLRGIEAQFLALQGFNEVIPHQLLTKTDFEKELEILICGLKID----- 667

623 ELICGLKID-----VNDMKVNTLKHCTPDSNIVKWFVKAFFPDEERRARLLQFVGGSSR 667

3915 ELLISGLPEIDCKLSDIQLCAYPLIDDLKANTVTSYTAGSPVIRWVWVKAFFKEDMARFLQFVGTGSK 3974

668 RARLLQFVTGSSRVPLOGFKALQAGAPRLFTTHQIDACTNNLPKHAHTCFNRIDIPPYSYKLYKLLTAIE 727

3975 MARFLQFVTGSSRVPLOGFKALQAGAPRLFTTHQIDACTNNLPKHAHTCFNRIDIPPYSYKLYKLLTAIE 727

728 YEKLYEKLTAIEBTC---GFA 746

4035 KEQLQERLLAIHASEGFGFA 4056

RESULT 14

H96625

hypothetical protein YDR457w - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 23-Mar-2001

C:Accession: S69625

R:Dietrich, F.S.

submitted to the EMBL Data Library, August 1995

A:Description: The sequence of S. cerevisiae cosmid 9410, 8035, 8166, and 9787.

A:Reference number: S69554

A:Accession: S69625

A:Molecule type: DNA

A:Residues: 1-3268 <DIE>

A:Cross-references: EMBL:U33050; NID:g927726; PID:AA864910.1; PID:g927738; MIPS:YDR457w

C:Genetics:

A:Gene: SGD:TOM1

A:Cross-references: SGD:S0002865; MIPS:YDR457w

A:Map position: 4R

Query Match 19.5%; Score 789; DB 2; Length 3268;

Best Local Similarity 45.7%; Pred. No. 5.1e-48;

Matches 164; Conservative 59; Mismatches 132; Indels 4; Gaps

384 SQQPOAGCHIEVSREIPEESYRQVWKMPPKDLWKMLIKFRGEGLDYGGVAREWL 442

2905 NOERPK---LPITVRRQVFLDSYALFFKTNDSIKNSKLEITPKGSGVDAGVTREWY 2961

443 YLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHLSYFHFVGRINGMAVFHGHYIDG 502

2962 QVLSRQMFNPYALFLPVPDSKTTFFHNPRTSGINPEHLSFFKFGIMIGKAIHQCELDLC 3021

QY 503 GFTLPFYKOLLGKSTLLDMELVDPDLHNSLVWILLENDITGVLDHTPCVEHNAYGEIIQH 562  
 Db 3022 HFSRVYKILGRPVSLKDMESLDDPYKSLVWILLENDITIIETFSVETDDYGEHKVI 3081  
 QY 563 ELKPNKGKSIPIVNEENKKEYRLVYVNRFLRGIEAQLALQKGFNEVPIQHLLKTFDEKEL 622  
 Db 3082 NLIEGGKDIIVTEANKQDYKVKVVEYKLTQTSKEQMDNLFVGLYALISKDLITIFDEQEL 3141  
 QY 623 ELIIICGLKIDVNDKVNTRLKHCTPDNSNIYKWFKAVEFDEERARLLQFVTGSSRVP 682  
 Db 3142 ELIIISGLPIDVDDKNNNTVYNYTATCKEVSYPFRAVRSFDEERAKLLQFVTGTSKVP 3201  
 QY 683 LQGFALOGAAGPRFTIHOIDACTNNLPKHAHTCFNRIDIPYSEYKLYEKLITAJEE 741  
 Db 3202 LNGFKELSGVGVCKFSIHRFGSSERLPSSHTCFNQLNLPYSEYETLRGSLLLAINE 3260

## RESULT 15

T20274  
 hypothetical protein F45H7.6 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Aug-2000  
 C:Accession: T20274; T22257  
 R:Percy, C.  
 submitted to the EMBL Data Library, November 1994  
 A:Reference number: Z19245  
 A:Accession: T20274  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-889 <W11>  
 A:Cross-references: EMBL:Z46793; PIDN:CAA86773.1; GSPDB:GN00021; CESP:F45H7.6  
 A:Experimental source: clone C56G7  
 R:Percy, C.  
 submitted to the EMBL Data Library, June 1994  
 A:Reference number: Z19538  
 A:Accession: T22257  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-889 <W12>  
 A:Cross-references: EMBL:Z34800; PIDN:CAA84325.1; GSPDB:GN00021; CESP:F45H7.6  
 A:Experimental source: clone F45H7  
 C:Genetics:  
 A:Gene: CESP:F45H7.6  
 A:Map position: 3  
 A:Introns: 30/3; 54/1; 118/1; 185/2; 213/1; 260/3; 288/2; 353/2; 454/1; 559/2; 642/2; 74/2  
 A:Superfamily: WW repeat homology  
 F:229-266/Domain: WW repeat homology <WW1>  
 F:372-408/Domain: WW repeat homology <WW2>

Query Match 17.6%; Score 709.5; DB 2; Length 889;  
 Best Local Similarity 27.2%; Pred. No. 4.1e-43;  
 Matches 194; Conservative 126; Mismatches 271; Indels 123; Gaps 19;  
 QY 131 QIVVSLQSRDRIGTGQVVDGCSRLPDNDLPDQWEERRTASGRIOVLNHIHTRITOWERPTR 190  
 Db 203 EIVFSIEPKTSKSPADIMEIASSQTPESHKTYLDKKGKFWVNHVTKETWTKPDT 262  
 QY 191 PASEYSSGRPLSCFVDENTPIGTNGATCGQSDPR---LAERVRSSQRHR----- 239  
 Db 263 LNNNHIEPETPVHKRLSD-----RSASPRNSFITPRRTITVRSAGCPKSDLI 309  
 QY 240 NYMSRTHLTPDLPEGVEQRTQGVFLHTQTGVSTGVDHDPVRDLT----- 289  
 Db 310 QFFQDEKFTAL-----YENQDAMQ-----IYNCSVVRHAIHQKDLDPSPKFNQPL 359  
 QY -290 -----NINCEELGPIPPQWEIRNTATGRVYVDHNNRTTQFTDPRL-----SANLHV 337  
 Db 360 FVRVFNLPADITQPLPSGWEC-ITMNNRTVFLNHANKETSFYDPRIRRPETKTSRRGRSV 418  
 QY 338 LNRQNQL-KDQQQQVSVSLCPDDTEC-----LTVPRYKEDL 372  
 Db 419 PSRSSTAHKGIDHALISKCEDLRKIAQDNFPQIAERISKXLMILIERFGGLAVASLANDL 478

QY 373 -----VQKL-----KIIRQELSSQQQPOAGHCRI---EVSREEIFEESYRQVM 411  
 Db 479 DITLALSMDSNTEKLAGEDNINMYEDMKKEKLGKSPSRLCWKYSRDLDDAFRIIL 538  
 QY 412 KMHPKOLWK-RLMIKPERGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQIN 470  
 Db 539 NVDPFVLKKSRLHIREGELALDYGGLSREFFTLISRELPHPKNGYFEYEGND-YHLQLR 597  
 QY 471 PDSAVNPEHLSYFHFVGRINGMAVFGHYIDGGFTLPPFYKQLLGKSIITDDMELVDPDLH 530  
 Db 598 PRGCETEKEKKWILICORVLALAVIHRCTYDVFTNVFYKSLQKRPVTLMDFKESDAEFY 657  
 QY 531 NSLVWILLENDITGVLDHTPCVEHNAYG-----EIIQHELKPNKGSIPVNE 575  
 Db 658 KSMNWILLENDVVD-LEMSFYSSWNGKVSDDLTLSSIVSVALQLAQELVPGGESQMVTE 716  
 QY 576 ENKKEYVRLVYVNRFLRGIEAQLALQKGFNEVPIQHLLKTFDEKELEIIICGLGKIDVN 635  
 Db 717 ANKAEPIDLMCKOKKAIRGVEKPLEILLTSPNQILNDNLNLSLESSDLKRLSGSLDLN 776  
 QY 636 DWKVNTRLKHCTPDNSNI-VKWFKAVEFDEERARLLQFVTGSSRVPLOGFKALOGAAG 694  
 Db 777 DWRTNIYKGGYSDCHIVVEWFWFWEVETWNTQERFDLLLFTVGSSVPPFEGFSAALRGNEE 836  
 QY 695 PRLEFTHIQIDACTNNLPKHAHTCFNRIDIPYSEYKLYEKLITAJEECTGFAVE 748  
 Db 837 ISKFCIEKMGDAI-SFPRHTCFNRLQLPSYNTKQKLSKQLQQAIVNGMSYSIE 889

Search completed: September 21, 2004, 07:48:30  
 Job time : 35.0014 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:29:24 ; Search time 17.2889 Seconds  
(without alignments)  
2252.800 Million cell updates/sec

Title: US-10-009-945-4

Perfect score: 4038

Sequence: 1 MSNFCRRRNGPVKRLTLVLC.....EKLVEKLLTAIETCTGPAVE 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4031	99.8	748	1	SUF2 HUMAN
2	3015.5	74.7	757	1	Q9hau4 homo sapien
3	2937.5	72.7	731	1	Q9p2n2 xenopus lae
4	2491.5	61.7	619	1	Q8cun6 mus musculus
5	1585.5	39.3	767	1	Q92462 schizosacch
6	1484	36.8	809	1	Q92462 schizosacch
7	1358	33.6	887	1	P46935 mus musculus
8	1355.5	33.6	887	1	Q62940 rattus norv
9	1354.5	33.5	1000	1	P46934 homo sapien
10	1336	33.1	864	1	Q8c863 mus musculus
11	1313.5	32.5	903	1	Q96j02 homo sapien
12	1298.5	32.2	870	1	Q00308 homo sapien
13	1297	32.1	922	1	Q9h0m0 homo sapien
14	1286.5	31.9	870	1	Q9dbh0 mus musculus
15	1281	31.7	918	1	Q8bz23 mus musculus
16	626	15.5	310	1	P51593 rattus norv
17	539	13.3	875	1	Q05086 homo sapien
18	484.5	12.0	885	1	Q08759 mus musculus
19	460.5	11.4	892	1	P40985 saccharomyc
20	445.5	11.0	1050	1	Q15034 homo sapien
21	404	10.0	1992	1	Q14669 homo sapien
22	342	8.5	910	1	P53119 saccharomyc
23	286.5	7.1	1620	1	Q9ult8 homo sapien
24	281	7.0	920	1	Q62671 rattus norv
25	280	6.9	2799	1	Q95071 homo sapien
26	278.5	6.9	1483	1	P3202 saccharomyc
27	262.5	6.5	1647	1	Q10435 schizosacch
28	224.5	5.6	2895	1	P51592 drosophila
29	207	5.1	472	1	P46938 mus musculus
30	161.5	4.0	1277	1	Q88382 rattus norv
31	159	3.9	1275	1	Q9wq1 mus musculus
32	158.5	3.9	1455	1	Q66ul8 homo sapien
33	158	3.9	446	1	P46936 gallus gall

34	154.5	3.8	454	1	YAP1_HUMAN	P46937 homo sapien
35	129	3.2	608	1	SAV_DROME	Q9vcr6 drosophila
36	127.5	3.2	1530	1	RIM2_MOUSE	Q9eqz7 mus musculus
37	125	3.1	386	1	SAV1_MOUSE	Q8ved2 mus musculus
38	122.5	3.0	672	1	KPCA_HUMAN	P17252 homo sapien
39	122.5	3.0	672	1	KPCA_RABIT	P10102 oryctolagus
40	122.5	3.0	672	1	KPCA_RAT	P05696 rattus norv
41	122	3.0	637	1	GYPI_YEAST	Q08484 saccharomyc
42	121	3.0	1188	1	RIM2_HUMAN	Q9uq26 homo sapien
43	120.5	3.0	383	1	SAV1_HUMAN	Q9h4b6 homo sapien
44	120.5	3.0	1813	1	UN13_CAEEL	P27715 caenorhabdi
45	120	3.0	1555	1	RIM2_RAT	Q9jisl rattus norv

#### ALIGNMENTS

##### RESULT 1

ID	SUF2 HUMAN	STANDARD;	PRT;	748 AA.
AC	Q9HAF4; Q9H260;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Smad ubiquitination regulatory factor 2 (EC 6.3.2.-) (Ubiquitin--protein ligase SMURF2) (Smad-specific E3 ubiquitin ligase 2)			
DE	(SMURF2).			
GN	SMURF2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND MUTAGENESIS OF 251-PRO--VAL-284 AND 297-GLY--LEU-330.			
RX	PubMed=11163210;			
RA	Kavsek P., Rasmussen R.K., Causing C.G., Bonni S., Zhu H.,			
RA	Thomsen G.H., Wrana J.L.;			
RT	"Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF-beta receptor for degradation.";			
RT	Mol. Cell 6:1365-1375(2000).			
RL	[2]			
RN	SEQUENCE FROM N.A., AND MUTAGENESIS OF 251-PRO--VAL-284 AND CYS-716.			
RP	MEDLINE=20538422; PubMed=11016919;			
RX	Lin X., Liang M., Feng X.-H.;			
RA	"Smurf2 is a ubiquitin E3 ligase mediating proteasome-dependent degradation of Smad2 in transforming growth factor-beta signaling.";			
RT	J. Biol. Chem. 275:36818-36822(2000).			
RL	[3]			
RN	SEQUENCE FROM N.A., AND MUTAGENESIS OF CYS-716.			
RP	MEDLINE=21107656; PubMed=1158580;			
RX	Zhang Y., Chang C., Gehling D.J., Hemmati-Brivanlou A., Derynck R.;			
RA	"Regulation of Smad degradation and activity by Smurf2, an E3 ubiquitin ligase.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 98:974-979(2001).			
CC	-/- FUNCTION: Interacts with SMAD1, SMAD2 and SMAD7 in order to trigger their ubiquitination and proteasome-dependent degradation. Enhances the inhibitory activity of SMAD7 and reduces the transcriptional activity of SMAD2. Coexpression of SMURF2 with SMAD1 results in considerable decrease in steady-state level of SMAD1 protein and a smaller decrease of SMAD2 level.			
CC	-/- SUBUNIT: Interacts with SMAD1, SMAD2, SMAD3, SMAD6 and SMAD7 but not SMAD4.			
CC	-/- SUBCELLULAR LOCATION: Nuclear. Cytoplasmic in the presence of SMAD7.			
CC	-/- TISSUE SPECIFICITY: Widely expressed.			
CC	-/- DOMAIN: The second and third WW domains are responsible for interaction with R-SMAD (SMAD1, SMAD2 and SMAD3).			
CC	-/- SIMILARITY: Contains 1 C2 domain.			
CC	-/- SIMILARITY: Contains 3 WW domains.			
CC	-/- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase domain.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [licenses@isb-sib.ch](mailto:licenses@isb-sib.ch)).  
 CC -----  
 DR ENBL; AF310676; AAC45422.1; -;  
 DR ENBL; AF310463; AAC25641.1; -;  
 DR ENBL; AY014180; AAC50421.1; -;  
 DR HSP; Q13526; IPIN.  
 DR MM; 605532; -;  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; NAS.  
 DR GO; GO:0016481; P:negative regulation of transcription; NAS.  
 DR GO; GO:0017015; P:regulation of TGFbeta receptor signaling pa...; NAS.  
 DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; NAS.  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR008973; C2\_CaLB.  
 DR InterPro; IPR000569; HECT domain.  
 DR InterPro; IPR001202; WW\_Rep5\_WWP.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00632; HECT; 1.  
 DR Pfam; PF00337; WW; 3.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00119; HECTC; 1.  
 DR SMART; SM00456; WW; 3.  
 DR PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
 DR PROSITE; PS00004; C2\_DOMAIN\_2; 1.  
 DR PROSITE; PS0237; HECT; 1.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 1.  
 DR PROSITE; PS00020; WW\_DOMAIN\_2; 3.  
 DR Ubl conjugation pathway; Ligase; Repeat; Nuclear protein.  
 FW DOMAIN 1 98 C2 DOMAIN.  
 FT DOMAIN 157 190 WW 1.  
 FT DOMAIN 251 284 WW 2.  
 FT DOMAIN 297 330 WW 3.  
 FT DOMAIN 414 748 HECT.  
 FT MUTAGEN 251 284 MISSING: ABOLISHES INTERACTION WITH  
 FT MUTAGEN 297 330 SMAD2 AND SMAD7.  
 FT MUTAGEN 716 716 MISSING: ABOLISHES INTERACTION WITH  
 FT MUTAGEN 716 716 C-2A: LOSS OF ABILITY TO UBIQUITINATE  
 FT MUTAGEN 716 716 C-2G: LOSS OF ABILITY (LOSS OF ABILITY  
 FT TO UBIQUITINATE SMAD1 AND SMAD2 AND NO  
 FT DOWN-REGULATION OF SMAD1 AND SMAD2  
 FT PROTEIN LEVELS).  
 FT CONFLICT 6 6 G -> R (IN REF. 2).  
 SQ SEQUENCE 748 AA; 86195 MW; 3042B443A3755762 CRC64;  
 Query Match 99.8; Score 4031; DB 1; Length 748;  
 Best Local Similarity 99.9; Pred. No. 6.ee-280;  
 Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 1 MSNPGRRNGPVKLRITLVCAKLVKDFRLLPDPFAKVVVDGSGGCHSTDTVKNLDPK 60  
 1 MSNPGRRNGPVKLRITLVCAKLVKDFRLLPDPFAKVVVDGSGGCHSTDTVKNLDPK 60  
 61 WNHQYDLYIGKSDSVTISVNNHKKIHKQAGFLGCVRLLSNAINLKTGYQRLDLCKL 120  
 61 WNHQYDLYIGKSDSVTISVNNHKKIHKQAGFLGCVRLLSNAINLKTGYQRLDLCKL 120  
 121 GPNDNDTVRGQIVVSLQSRDRIGTGGVQVDCSLFNDLPDQWEERTASGRQYLNHIT 180  
 121 GPNDNDTVRGQIVVSLQSRDRIGTGGVQVDCSLFNDLPDQWEERTASGRQYLNHIT 180  
 181 RTTOWERPTTPASYSYSGRPLSCFVDENTPIISGTNGATCGSSDPLAERRVRSQRHN 240  
 181 RTTOWERPTTPASYSYSGRPLSCFVDENTPIISGTNGATCGSSDPLAERRVRSQRHN 240  
 241 YMSRTHLHTPDPEGYEQRTTQGGQVYFLHTGTGVSTWHDPRVPRDLNSNINCEELGPLP 300

Db 241 YMSRTHLHTPDPEGYEQRTTQGGQVYFLHTGTGVSTWHDPRVPRDLNSNINCEELGPLP 300  
 Qy 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQOLKQDQOQVYVSLCPDDT 360  
 Db 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQOLKQDQOQVYVSLCPDDT 360  
 Qy 361 ECLTVPRYKRDVLVOKLILRQELSSQOQPOQAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420  
 Db 361 ECLTVPRYKRDVLVOKLILRQELSSQOQPOQAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420  
 Qy 421 RLMIKFRGEGLDYGVARWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480  
 Db 421 RLMIKFRGEGLDYGVARWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480  
 Qy 481 SYFHFVGRINGMAVFHGHYIDGGFTLPFYKQLLGKSTLDDMELVDPDLHNSLVWILEND 540  
 Db 481 SYFHFVGRINGMAVFHGHYIDGGFTLPFYKQLLGKSTLDDMELVDPDLHNSLVWILEND 540  
 Qy 541 ITGVLDHTFCVEHNAYGEIIQHLEKPNKKSIPYNEENKKEYVRLYVNWRLRGLGIEAQFLA 600  
 Db 541 ITGVLDHTFCVEHNAYGEIIQHLEKPNKKSIPYNEENKKEYVRLYVNWRLRGLGIEAQFLA 600  
 Qy 601 LQKGFNEVIPQHLKLTFTDEKELELIICGLGKIDVNDWKVNTLKHCTPDSNIVKWFKAV 660  
 Db 601 LQKGFNEVIPQHLKLTFTDEKELELIICGLGKIDVNDWKVNTLKHCTPDSNIVKWFKAV 660  
 Qy 661 EFDERRARLLQFVTGSSRVPLQGFALOGAAGPRLTIHQIDACTNNLPKHAHTCENRI 720  
 Db 661 EFDERRARLLQFVTGSSRVPLQGFALOGAAGPRLTIHQIDACTNNLPKHAHTCENRI 720  
 Qy 721 DIPVESYEKLYEKLTAIBETCGFAVE 748  
 Db 721 DIPVESYEKLYEKLTAIBETCGFAVE 748  
 RESULT 2  
 SUF1\_HUMAN  
 ID SUF1\_HUMAN STANDARD; PRT; 757 AA.  
 AC Q9HCE7; O75853; Q9JUT8.  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--  
 DE protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)  
 DE (hSMURF1).  
 DE SMURF1 OR KIAA1625.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE OF 10-731 FROM N.A. (ISOFORM SHORT).  
 RX MEDLINE=9385348; PubMed=10458166;  
 RA Zhu H., Kavsak P., Abdollah S., Wrana J.L., Thomsen G.H.;  
 RT "A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic  
 RT pattern formation."  
 RL Nature 400:687-693(1999).  
 RN [2]  
 RP SEQUENCE OF 20-731 FROM N.A. (ISOFORM SHORT).  
 RA Stoneking T., Bauer C., O'Neal D.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM LONG).  
 RC TISSUE=Brain;  
 RX MEDLINE=20450683; PubMed=1097877;  
 RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes.  
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which  
 RT code for large proteins in vitro."  
 RL DNA Res. 7:273-281(2000).  
 CC -!- FUNCTION: Interacts with receptor-regulated SMADs specific for the  
 CC bmp pathway (SMAD1 AND SMAD5) in order to trigger their  
 CC ubiquitination and degradation and hence their inactivation.



(ACTIVIN/TGF-BETA) PATHWAY.

CC -1- DEVELOPMENTAL STAGE: EXPRESSED FROM THE EGG STAGE TO THE SWIMMING  
 CC TADPOLE, WITH MAXIMUM LEVELS OBSERVED IN THE STAGES FROM EGG TO  
 CC GASTRULA. AT GASTRULATION DISTRIBUTED UNIFORMLY IN EMBRYONIC  
 CC ECTODERM AND INVOLUTING MESODERM, AND EXPRESSION GRADUALLY  
 CC LOCALIZES TO THE NERVOUS SYSTEM, AT EARLY TADPOLE STAGES EXPRESSED  
 CC IN THE CNS, EYE, BRANCHIAL ARCHES, KIDNEY AND SOMITES.

CC -1- SIMILARITY: Contains 1 C2 domain.  
 CC -1- SIMILARITY: Contains 2 WW domains.  
 CC -1- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL, AF169310; AD52564.1; --  
 CC GO: GO:0005622; C:intracellular; TAS.  
 CC GO: GO:0002111; F:protein degradation tagging activity; IDA.  
 CC GO: GO:0004842; F:ubiquitin-protein ligase activity; IDA.  
 CC GO: GO:0030154; P:cell differentiation; IDA.  
 CC GO: GO:0007398; P:ectoderm development; TAS.  
 CC GO: GO:0030514; P:negative regulation of BMP signaling pathway; TAS.  
 CC GO: GO:0006512; P:ubiquitin cycle; IDA.

CC InterPro: IPR000008; C2.  
 CC InterPro: IPR008973; C2\_CaLB.  
 CC InterPro: IPR000569; HECT domain.  
 CC InterPro: IPR001202; WW\_Rsp\_WWP.  
 CC Pfam: PF00168; C2; 1.  
 CC Pfam: PF00632; HECT; 1.  
 CC Pfam: PF00397; WW; 2.  
 CC SMART: SM00239; C2; 1.  
 CC SMART: SM00119; HECT; 1.  
 CC SMART: SM00456; WW; 2.  
 CC PROSITE: PS00499; C2\_DOMAIN\_1; 1.  
 CC PROSITE: PS00004; C2\_DOMAIN\_2; 1.  
 CC PROSITE: PS02027; HECT; 1.  
 CC PROSITE: PS01159; WW\_DOMAIN\_1; 1.  
 CC PROSITE: PS00020; WW\_DOMAIN\_2; 2.  
 CC Us1 conjugation pathway; Ligase; Repeat.

CC FT DOMAIN 1 99 C2 DOMAIN.  
 CC FT DOMAIN 233 266 WW 1.  
 CC FT DOMAIN 279 312 WW 2.  
 CC FT DOMAIN 394 731 HECT.  
 CC SEQUENCE 731 AA; 83259 MW; 3CE88E512A42CE2C CRC64;

Query Match 72.7%; Score 2937.5; DB 1; Length 731;  
 Best Local Similarity 72.7%; Pred. No. 7.1e-202;  
 Matches 558; Conservative 66; Mismatches 85; Indels 57; Gaps 9;

QY 1 MSNPGRRNG-PYKRLTVLCANLYKKOFFRDPDPFAKVVDSGSOCHSTDVKNLDP 59  
 DB 1 MSNVVTRGSSIRVRLTVLCANLYKKOFFRDPDPFAKVVDSGSOCHSTDVKNLDP 50  
 QY 60 KWNCHYDLYGKSDSVTISVWNNKKHKQAGFLGCVRLLSNAINRLKDTGYQRLDCK 119  
 DB 61 KWNCHYDLYGKSDSVTISVWNNKKHKQAGFLGCVRLLSNAINRLKDTGYQRLDCK 120  
 QY 120 LGPNDNDTVRGQIVSLQSDRGTGQGVVDCSRLFDNDLPDQWEERRTASGRIQVNIH 179  
 DB 121 LNPTDNDVARGQIVSLQSDRGTGQGVVDCSRLFDNDLPDQWEERRTASGRIQVNIH 162  
 QY 180 TRTQWERTRPASEVSPGRPLSCFVDENTPIS---GTNGATCGSSDPRLAERVRSQ 236  
 DB 163 -----LEDTGPRPLSCFMDERPYTDGPGAGGGPGRLVESPGQEQRLQAQ 209  
 QY 237 RHRNMYGTHLHTP-----PDLPEGVEQRTQGGVYFLHTGTGVTSHDPRVPD 287  
 DB 210 RVRGPEVREHVQTPQNRSHGFSQDLPEGEYEQRTTQGGVYFLHTGTGVTSHDPRVPD 269

QY 288 LSNINCEELGPLEPGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLNRQNLKQ 347  
 DB 270 LNSVNCDDGLSLPAGMEVRTTVSGRIYFVDHNNRTTQFTDPR-----LHIIHQSLKSP 325  
 QY 348 QQ-----QQVYSLCPDDTECLTVPRYKRDVLQKILRQELSQOQOAGHCRVRSREIF 403  
 DB 326 NHAIPVQSGSL--EDGDEPPAQRYERDLVQKLVLRHLSLQPOAGHCRVRSREIF 383  
 QY 404 BESYRQWKRPKDLKRLMKIKRGEGIDYGVAREWYLLSHEMLNPYGLFOYSRDD 463  
 DB 384 BESYRQIMKRPKDLKRLMKIKRGEGIDYGVAREWYLLSHEMLNPYGLFOYSTDN 443  
 QY 464 IYTLQINPDGAVNPEHLSYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGKSIITDDME 523  
 DB 444 IYTLQINPDSINSYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGKSIITDDME 503  
 QY 524 LVDPDLHNSLVILENDITGVLDHTPCVHNNAVEIIQHELKPKNGKSIIPVNEENKKEYVR 593  
 DB 504 SVDPDLHNSLVILENDITGVLDHTPCVHNNAVEIIQHELKPKNGKSIIPVNEENKKEYVR 563  
 QY 584 LYVNWRLGIEAQFLAQKGFNEVIPHLLKTFDEKELELIICGLKIDVNDKVNTRL 643  
 DB 564 LYVNWRLGIEAQFLAQKGFNEVIPHLLKTFDEKELELIICGLKIDVNDKVNTRL 623  
 QY 644 KHCTPDSNIVKWPWKAYEFFDEERRARLLOFVTGSSRVPLQGGKALQ---GAAGPLFTI 700  
 DB 624 KHCLANSNIVQWFWQAVESFDEERRARLLOFVTGSTRVPLQGGKALQGGSTGAAGPLFTI 683  
 QY 701 HQIDACTNNLPKHAHTCFNRDIPPEYSEKLYEKLTAETCTGFAVE 748  
 DB 684 HLIDANDTNLPKHAHTCFNRDIPPEYSEKLYEKLTAETCTGFAVE 731

RESULT 4  
 SUF1\_MOUSE  
 ID SUF1\_MOUSE STANDARD; PRT; 619 AA.  
 AC Q9CUN6;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Smad ubiquitination regulatory factor 1 (EC 6.3.2.-) (Ubiquitin--  
 DE protein ligase SMURF1) (Smad-specific E3 ubiquitin ligase 1)  
 DE (Fragment).  
 GN SMURF1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaide I., Osato R., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojibori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chochia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grammond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,

RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
CC -!- FUNCTION: Interacts with receptor-regulated SMADs specific for the  
CC bmp pathway (SMAD1 AND SMAD5) in order to trigger their  
CC ubiquitination and degradation and hence their inactivation (By  
CC similarity).  
CC -!- SIMILARITY: Contains 1 C2 domain.  
CC -!- SIMILARITY: Contains 2 WW domains.  
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
CC domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AK015264; BAB29770.2; AUT\_INIT.  
DR HSSP; Q13526; IPIN.  
DR MGD; MG1:1923038; 4930431E10R.k.  
DR GO; GO:0005622; C:intracellular; ISS.  
DR GO; GO:0000211; F:protein degradation tagging activity; ISS.  
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; ISS.  
DR GO; GO:0030154; P:cell differentiation; ISS.  
DR GO; GO:0007398; P:ectoderm development; ISS.  
DR GO; GO:0030514; P:negative regulation of BMP signaling pathway; ISS.  
DR GO; GO:0006512; P:ubiquitin cycle; ISS.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR000569; HECT domain.  
DR InterPro; IPR01202; WW\_RSP5\_WWP.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 2.  
DR SMART; SM00119; HECTc; 1.  
DR SMART; SM00456; WW; 2.  
DR PROSITE; PS00499; C2 DOMAIN 1; PARTIAL.  
DR PROSITE; PS00004; C2 DOMAIN 2; PARTIAL.  
DR PROSITE; PS00237; HECT; 1.  
DR PROSITE; PS01159; WW DOMAIN 1; 1.  
DR PROSITE; PS00020; WW DOMAIN 2; 2.  
KW Ub1 conjugation pathway; Ligase; Repeat.  
FT NON\_TER 1  
FT DOMAIN 125 158 WW 1.  
FT DOMAIN 171 204 WW 2.  
FT DOMAIN 285 619 HECT.  
FT DOMAIN 35 38 POLY-GLY.  
SQ SEQUENCE 619 AA; 70902 MW; 4CBE2F8624A7B525 CRC64;  
Query Match 61.7%; Score 2491.5; DB 1; Length 619;  
Best Local Similarity 72.0%; Pred. No. 3.8e-170;  
Matches 471; Conservative 59; Mismatches 75; Indels 49; Gaps 6;  
QY 109 DTGYORLDLKGPNNDTVRGQIVVLSQSDRDTGTGGQVDCSRLPDNDLPDGCWEERT 169  
DB 1 DTGYORLDLCLNPSDPTDAVRGQIVVLSQSDRDTGTGGQVDCSRLPDNDLPDGCWEERT 50  
QY 169 ASGRIOYLNHTRITQWERPRPASEYSSPGRLPSCFVDENPTISGNTGATCG----- 221  
DB 51 --GTVY-----EDSGPRPISCLNEEPAFTDGTGAAGGNCRFV 89  
QY 222 --QSSDPLAERRVRSQRHNYM----SRTHLHTPDPDLPGEYQRTTQGGQVYFLHTQTG 275  
DB 90 ESPSCNQRLVQLONPEVQPLQPPQNRPHGQSPELPEGEYQRTTQGGQVYFLHTQTG 149  
QY 276 VSTWHDPRVPRDLSNINCEELGPLPPGWEIRNTATGRVYFVDHNNRTTQFDPRLSANLH 335

Db 150 VSTWHDPRVPRDLSNINCEELGPLPPGWEIRNTATGRVYFVDHNNRTTQFDPRLSANLH 205  
QY 336 LVLNQRNQLKDDQQQ-QVYSLCPDDTTECLTVPRYKDLVQKLKILRQELSSQQQPOAGHCR 394  
Db 206 HINWQOQLKEPSQPLQNEGSVEDELPAPRYERDLVQKLKILRQELSSQQQPOAGHCR 265  
QY 395 IEVSREEIEFEESYRQVMKRPKDLWKRLMKFRGEEGLDYGGVAREWLYLLSHEMLNPPY 454  
Db 266 IEVSREEIEFEESYRQVMKRPKDLWKRLMKFRGEEGLDYGGVAREWLYLLSHEMLNPPY 395  
QY 455 GLFOYRSDDIYTLQINPDNAVNPSHLSYHFVGRINGMAVFGHYIDGGFTLPFYKQLLG 514  
Db 326 GLFOYSTDNITLTQINPDSSINPDHLSYHFVGRINGMAVFGHYIDGGFTLPFYKQLLG 385  
QY 515 KSITLDMELVDPDLHNSLWILENDITGVLDHTFCVHNAYGEIIQHELKPKNGKIPVN 574  
Db 386 KPQLSDLESVDPELKHSLWILENDITPVLDHTFCVHNAYGEIIQHELKPKNGRNPVT 445  
QY 575 EENKEEYVRLYNVWRFLRGIEAQFLAQKGFNEVIFQHLKTPDEKELELIICGLSKIDV 634  
Db 446 EENKEEYVRLYNVWRFLRGIEAQFLAQKGFNEVIFQHLKTPDEKELELIICGLSKIDV 505  
QY 635 NDWKVNTRLKCHCTPDNSNIVKFWKAVEFFDEERRAALLOFVTGSSRVPLQGGFKALQGAAG 694  
Db 506 NDWKSNTRLKCHCVADSNIVRWFQAVETDEERRAALLOFVTGSTRVPLOGFKALQGAAG 565  
QY 695 PRFLTTHQIDACTNNLPKHAHTCFNRIDIPPESEYKLYEKLITAEETCGFAVE 748  
Db 566 PRFLTTHLIDANTDNLPAKHAHTCFNRIDIPPESEYKLYEKLITAEVETCGFAVE 619  
RESULT 5  
PUBL SCHPO STANDARD; PRT; 767 AA.  
AC Q92462; O14454;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Ubiquitin--protein ligase pubi (EC 6.3.2.-).  
GN PUBI OR SPAC11G7.02.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96205868; PubMed=8635463;  
RA Nefsky B., Beach D.;  
RT "Publ acts as an E6-AP-like protein ubiquitin ligase in the  
RT degradation of cdc25";  
RL EMBO J. 15:1301-1312(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=J227;  
RA Saleki R., Jia Z., Karagiannis J., Young P.G.;  
RT "Tolerance of low pH in Schizosaccharomycetes pombe requires a  
RT functioning publ ubiquitin ligase";  
RL Mol. Gen. Genet. 254:520-528(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX STRAIN=972;  
RA MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,

RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grynponez B.,  
 RA Waltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Garzel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Babely K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Szpakowski G.V., Ussery D., Barrell B.G., Nurse P.,  
 RA "the genome sequence of Schizosaccharomyces pombe.",  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: REGULATES UBIQUITINATION OF CDC25.  
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +  
 CC diphosphate + protein N-ubiquityllysine.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thioester formation.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 3 WW domains.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.  
 CC -!- SIMILARITY: STRONG, TO YEAST RSP5.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; Y07592; CAA68867.1; -;  
 DR EMBL; U66716; AAB07514.1; -;  
 DR EMBL; Z99161; CAB16207.1; -;  
 DR EMBL; U62795; AAB63350.1; -;  
 DR PIR; S66562; S66562.  
 DR PIR; T37545; T37545.  
 DR HSP; Q13526; IPIF.  
 DR GeneDB Spombe; SPAC11G7.02; -;  
 DR InterPro; IPR000008; C2.  
 DR InterPro; IPR008973; C2 CalB.  
 DR InterPro; IPR000569; HECT\_domain.  
 DR InterPro; IPR002349; WW\_domain.  
 DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00632; HECT; 1.  
 DR Pfam; PF00397; WW; 3.  
 DR PRINTS; PR00403; WWDOMAIN.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00119; HECTC; 1.  
 DR SMART; SM00456; WW; 3.  
 DR PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
 DR PROSITE; PS50004; C2\_DOMAIN\_2; 1.  
 DR PROSITE; PS50237; HECT; 1.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 3.  
 DR PROSITE; PS50020; WW\_DOMAIN\_2; 3.  
 KW Ub1 conjugation pathway; Ligase; Repeat.  
 FT DOMAIN 17 32 C2 DOMAIN.  
 FT DOMAIN 211 236 WW 1.  
 FT DOMAIN 242 247 POLY-ALA.  
 FT DOMAIN 294 319 WW 2.  
 FT DOMAIN 351 376 WW 3.  
 FT DOMAIN 463 767 HECT.  
 FT BINDING 735 735 UBIQUITIN (BY SIMILARITY).  
 FT CONFLICT 163 163 Q -> K (IN REF. 1).  
 FT CONFLICT 609 609 MISSING (IN REF. 1).  
 FT CONFLICT 661 661 T -> K (IN REF. 1).

SQ SEQUENCE 767 AA; 87267 MW; F1455A155EB9ACF7 CRC64;  
 Query Match 39.38; Score 1585.5; DB 1; Length 767;  
 Best Local Similarity 41.3%; Pred. NO. 2.4e-105;  
 Matches 340; Conservative 121; Mismatches 231; Indels 131; Gaps 17;  
 QY 1 MSNFGRRNGPVLRLTLVLCANLKKXDFRFLPDPFAKVVVDGSGGCHSDTIVKNILDPK 60  
 DB 1 MSNSAQSR---RIRVIIVAADGLYKEDVFRFPDPFAVLVDGE-QTHHTTAIKTLNPNY 55  
 QY 61 WNYHYDIYIGKSDVTISVWNKHKKQAGFLGCVRLISNAINRLKDTGYORL--DLC 118  
 DB 56 WNETFEVNTDNTSIAQVFDQKKF-KKGGGFLGVNLRVGDVLDLAIGDEMLTRDLK 114  
 QY 119 KLGNDDNDTVRGQIWSLQ----- 137  
 DB 115 K--SNEVTVHGKIILNLSITTAQSTLQVPSAASGARTQTSITNDPQSSQSSSVSRNPA 172  
 QY 138 -----SRDRIGTGGQVDCSRLFDN-----DLPGWERTASGRIOYLHHTITTTQ 184  
 DB 173 SSRAGSPTRONAPASPASSSEPTFSFEDQYGLPFGWERTDNLGRTYVVDHTRSTT 232  
 QY 185 WERTP-----RPASEYSSPGRLSCFVDENTPISTNGATCGOSSDPRLAERRVRSORH 238  
 DB 233 WIRNLSVAGAAAAELHSSAS--SANVTGVQPSSSNAA-----RETSASVLT--- 279  
 QY 239 RNYMSRTHLHTPDLPEGYEQRTTQQGVYFLHTQTGVSTWHDPRVPRDLNIN----- 292  
 DB 280 ---SNATTAGSGELPGWEQRYTPGRPYFVDHNTRTTTTWDPRRQOYIRSYGGPNAT 335  
 QY 293 -----CEELGPLPGWEIRNTATGRVYFVDHNETTOFTDPRLSANLHLVNLNQNLKQD 347  
 DB 336 IQQPVQQLGELPGWEMRLNTARVYFVDHNTKTITTTWDPRLPSSL-----DQ 384  
 QY 348 QQQQVSLCPDDETCLETVPRYKRLVQKLKILRQLSQ--QFQAGHCRIEVSREEIFE 405  
 DB 385 -----NVPQYKRDREKLIVF---LSQPALHPLPGQCHIKVARNHIFED 425  
 QY 406 SYRQVMQRPKDLWKRLMKIFRGEGLDYGVAEWLYLSHEMLNPYGLFOYSRDDIY 465  
 DB 426 SYABIMRQSATDLKKRLMKIFDGEDGLDYGELGREYFLLSHEMFNPFYCLFEYSSVDNY 485  
 QY 466 TLQINPDSAVNPEHLSYFHFVGRIMGMAVPHGVHDGFTLPFFYKOLLGKSIITLDDMELV 525  
 DB 486 TLQINPHSGINPEHLNFKYFVIGVIGLAI FHRFVDAFFVYFVKMLQKKVTLQDMES 545  
 QY 526 PDLHNSLVITLNDITGVLDHTFCVEHNAYGEIIQHELKPNKSIIPVNEENKKEYVRLY 585  
 DB 546 DAERYSLVWILDNDITGVLDLTPSVEDNCFGEVVTIDLPKNGRNIETVEENKREYVDLV 605  
 QY 586 VNRFLRGIEAQLALOKGFNEVIPHQLLTKTFDBKELELIICGLGKIDVMDVKNVRLKH 645  
 DB 606 TVWRIQRIEIQFNAFEGSELIPQLINLVDFERELELLIGISEIDMEDKWKHTDYS 665  
 QY 646 CTSDSNIVKFWKAVFDFEERRARLLQFVTGSRVLPQGFKALQGAAGPRLFTTHIDA 705  
 DB 666 YSENDQIIKFEWELMDWESNEKSRLLQFTTGTSTRIPVNGFKDLQSGDGRPKFTIERAGE 725  
 QY 706 CTNNLPKAHTCFNRIDIPPYESYEKLYEKLTLALETTCGFAVE 748  
 DB 726 -PNKLPKAHTCFNRLDLPFYTSKDLHDHKLDAVEETIGFGQE 767  
 RESULT 6  
 RSP5\_YEAST  
 ID RSP5\_YEAST STANDARD; PRT; 809 AA.  
 AC P39940;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Ubiquitin-protein ligase RSP5 (BC 6.3.2.-)  
 GN RSP5 OR NP11 OR MDPI OR YER125W OR SYGP-ORF41.  
 OS Saccharomyces cerevisiae (Baker's yeast).





[illegible]



	967	967	UBIQUITIN (BY SIMILARITY).
BINDING	1000 AA;	114936 MW;	3728088E50C149CB CRC64;
SEQUENCE			

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sakai D., Shibata K., Shinagawa A.,  
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
RN [3]  
RN FUNCTION, AND INTERACTION WITH JUN AND JUNB.  
RP MEDLINE=21864584; PubMed=11828324;  
RX Medline=21864584; PubMed=11828324;  
RA Fang D., Elly C., Gao B., Fang N., Altman Y., Joazeiro C., Hunter T.,  
RA Copeland N., Jenkins N., Liu Y.C.;  
RA "Dysregulation of T lymphocyte function in itchy mice: a role for  
RT Itch in TH2 differentiation";  
RT Nat. Immunol. 3:281-287(2002).  
RL [4]  
RL INTERACTION WITH NOTCH1, AND MUTAGENESIS OF CVS-832.  
RP MEDLINE=20549573; PubMed=10940313;  
RX Medline=20549573; PubMed=10940313;  
RA Qiu L., Joazeiro C., Fang N., Wang H.-Y., Elly C., Altman Y., Fang D.,  
RA Hunter T., Liu Y.-C.;  
RA "Recognition and ubiquitination of Notch by Itch, a hec-type E3  
RT ubiquitin ligase";  
RL J. Biol. Chem. 275:35734-35737(2000).  
RN [5]  
RN INTERACTION WITH OGNL.  
RP MEDLINE=21895815; PubMed=11792481;  
RX Medline=21895815; PubMed=11792481;  
RA Traweger A., Fang D., Liu Y.-C., Stelzhammer W., Krizbai I.A.,  
RA Fresser F., Bauer H.-C., Bauer H.;  
RA "The tight junction-specific protein occludin is a functional target  
RT of the E3 ubiquitin-protein ligase itch";  
RL J. Biol. Chem. 277:10201-10208(2002).  
CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
CC then directly transfers the ubiquitin to targeted substrates.  
CC Regulates the transcriptional activity of several transcription  
CC factors, and probably plays an important role in the regulation of  
CC immune response.  
CC -!- PATHWAY: Ubiquitin conjugation; third step.  
CC -!- SUBUNIT: Interacts via its WW domains with OCLN, NOTCH1, JUN and  
CC JUNB. Interacts with DRP1A and NFE2 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q8C863-1; Sequence=Displayed;  
CC Note=Major form;  
CC Name=2;  
CC IsoId=Q8C863-2; Sequence=VSP\_008452, VSP\_008453;  
CC -!- TISSUE SPECIFICITY: Widely expressed.  
CC -!- PTM: Phosphorylated on tyrosine residues (By similarity).  
CC -!- DISEASE: Defects in Itch are the cause of the itchy phenotype  
CC which is an inflammatory and immunological condition characterized  
CC by inflammation in the lung and stomach, hyperplasia in lymphoid  
CC and hematopoietic cells and constant itching in the skin.  
CC -!- SIMILARITY: Contains 1 C2 domain.  
CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
CC domain.  
CC -!- SIMILARITY: Contains 4 WW domains.  
CC -!- CAUTION: It is uncertain whether Met-1 or Met-11 is the initiator.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF037454; AAB99764.1; ALT\_INIT.  
CC EMBL; AK048303; BAC33298.1; -.  
DR HSSP; Q13526; 1PIN.  
DR MGD; MGI:1202301; Itch.  
DR InterPro; IPR000008; C2.  
DR InterPro; IPR008973; C2\_CaLB.  
DR InterPro; IPR000569; HECT domain.  
DR InterPro; IPR002349; WW.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00168; C2; 1.  
DR Pfam; PF00632; HECT; 1.  
DR Pfam; PF00397; WW; 4.  
DR PRINTS; PRO0403; WWDOMAIN.  
DR SMART; SM00239; C2; 1.  
DR SMART; SM00119; HECTC; 1.  
DR SMART; SM00456; WW; 4.  
DR PROSITE; PS00499; C2 DOMAIN 1; FALSE\_NEG.  
DR PROSITE; PS00004; C2 DOMAIN 2; 1.  
DR PROSITE; PS0237; HECT; 1.  
DR PROSITE; PS01159; WW DOMAIN 1; 4.  
DR PROSITE; PS00020; WW DOMAIN 2; 4.  
KW Ub1 conjugation pathway; Ligase; Nuclear protein; Repeat;  
KW Phosphorylation; Alternative splicing.  
FT DOMAIN 5 99  
FT DOVAIN 287 320 WW 1.  
FT DOVAIN 319 352 WW 2.  
FT DOVAIN 399 432 WW 3.  
FT DOVAIN 439 472 WW 4.  
FT DOVAIN 530 864 HECT.  
FT DOVAIN 218 223 POLY-PRO.  
FT VARSPLIC 742 759 LLCQMQLDNDWQHAI -> MNFYLLKHTSKYSRYLF  
FT (in isoform 2).  
FT VARSPLIC 760 864 Missing (in isoform 2).  
FT MUTAGEN 832 832 /FTID=VSP\_008453.  
FT C->A: LOSS OF UBIQUITIN PROTEIN LIGASE  
FT ACTIVITY.  
SQ SEQUENCE 864 AA; 98993 MW; 905FDBE0A1EA7EA CRC64;  
Query Match 33.1%; Score 1336; DB 1; Length 864;  
Best Local Similarity 36.6%; Pred.No. 1.8e-87;  
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;  
QY 13 KRLTTLVCAK-NLVKDDFFRLDPFAKVVWDGSGQCHSTDTKNTLDPKWNQHYDLYIGK 71  
DB 19 QLQITVISAKLKENKWWFG-PSYVEVTVD--GQSKTEKCNNTNSPKWKQPLTVIVTP 75  
QY 72 SDSVTISVNNHKIHKK--QAGFLGCVRLLSNAINRLKDTGQRLDLCKLGNPDNDTVR 129  
DB 76 TSKLCFRVWSHQTLKSDVLLTAGLDIYETLKSNMKLEEV---VMTLQLVGCKEPTETM 132  
QY 130 GQIVV---SLOSRRDRTGGGVV-----DCSRLFDNDLPD-----GWEE 165  
DB 133 GDLSCVLDGLQVEAEVNTGETSSESTQDDDCRTRDTRVSTNGSEPEVAASENK 192  
QY 166 RRTASGRIQYLNHTRTQWERPTRPASEVSSGPRLSCFVDENT-----PIS 213  
DB 193 RANGNNSPLSNGGFKFSRPRSRPP--PTPRPASVNGSPSTNSDGSSTGSLPPT 250  
QY 214 GTN-----CATCG-----QSSDPLAERRVRSRHRNYM--- 242  
DB 251 NTNVTSTSEGATSGLIPLITISGSGRPLNTVTSQAPLPGEQV--DGRGRVYVDHV 309  
QY 243 -SRTHLTPPLPEGYEQRTTQQQVYFLHTQTGVSTWHDPRV----- 285  
DB 310 EKRTTWDRPELPFGWERRVDNMGRIVYVDHFTTTTQWRFTLESVRNVEQWQLQSOLQ 369  
QY 286 -----RDL-----SNINCEELGPLPGWEIRNTATGEVYVDHNNETTFD 327  
DB 370 GAOQFNORFIYGNQDLFATSQNKEDFLGPLPGWEKRTDSNGRVYFVHNHTRITQWED 429

QY 328 PRLSANLHLVNLQNLKQDQOQVSLCP---DTECLTV---PR-----Y 368  
 Db 430 PRSQOQ---LNEKLPGEWNRFTVDGIPYFVDHNRATTVIDPRTGKSALDNGPQIAY 485  
 QY 369 KRDVLQKILR---QELSQOQFQAGHCRIEYSREEIFEESYRQVMKVRPKDLKRLMIK 425  
 Db 486 VRDFRAKVQYFRFCQQLAMPQ---HIKITVTRKTLFEDSFQQIMSFQDRLRLWVI 541  
 QY 426 FGEGLDGGVAREWVLLSHLEMLNPPYGLFQYSRDDIYTLQINPDVAVNPEHLSYHPF 485  
 Db 542 FGEGLDGGVAREWVLLSHLEMLNPPYGLFQYSRDDIYTLQINPDVAVNPEHLSYHPF 601  
 QY 486 VGRIMGVAFVGHYIDGGFTLFFYKQLLCKSTLDDMELVDPDLNSLVWILENDITGV- 544  
 Db 602 IGRFIAMALFHGKFDITGSLFYLKLPKLVGLKDLSDIDPEFFNSLIWKENNISCEG 661  
 QY 545 LDHTICVEINAYGEIIOHLEKNGKSIIPVNBENKKEYVRLYNWFLRGIEAQFIALQKG 604  
 Db 662 LEMYSVDKLEILGEIKSHDLKENGNNILVTEENKEEYIRMAVMAELSRGVEEQTAFFEG 721  
 QY 605 FNEVIPQHLKTFDEKELELICGLGKIDVNDKVNTRLKCHCTPDSNTVKVFWKAVEFFD 664  
 Db 722 FNEIIPQYLOVFDKLEVLGCGQEDLNDQWQHAHYRHYTRTSKQIMFWQVKEID 781  
 QY 665 EERRARLQFVTGSRVPLQGFQKALQGAAGPRLFTIHOIDACTNNLPKAHTCFNRIDIPP 724  
 Db 782 NEKNRLLQFVTGTCRLPVGGFADLMGNGPQKFCIEKVGK-ENWLPRSHTCFNRDLDP 840  
 QY 725 YESYEKLYEKLTAETEICGRAVE 748  
 Db 841 YKSYQLKEKLFLAETEETGFGQE 864

## RESULT 11

ITCH HUMAN  
 ID ITCH HUMAN STANDARD; PRT: 903 AA.  
 AC Q96J02; O43584; Q96F66; Q9BY75; Q9H451; Q9HAU5;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Itchy homolog E3 ubiquitin protein ligase (EC 6.3.2.-) (ITCH)  
 DE (Atrophin-1-interacting protein 4) (ALP4) (NFE2-associated polypeptide  
 DE 1) (NAPP1).  
 GN ITCH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH NFE2.  
 RC TISSUE=Leukemia;  
 RX MEDLINE=21218930; PubMed=11318614;  
 RA Yang-Peng T.L., Shen C.K.J.;  
 RT "Human ITCH is a co-regulator of the hematopoietic transcription  
 RT factor NF-E2".  
 RL Genomics 73:238-241(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RA Miyazaki K., Okamoto Y., Sakamoto M., Nakagawara A.;  
 RT "Homo sapiens mRNA for ubiquitin protein ligase Itch, complete cds.";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21638749; PubMed=11780052;  
 RX Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Clee C.M.,  
 RA Chapman J., Clamp M., Clark L.N., Clark S.Y., Clegg S., Clegg S.,  
 Clegg S., Cogley V.E., Collier R.E., Connor R., Corby N.R.,  
 Coulson A., Coville G.J., Deadman R., Dhani P., Dunn M.,  
 Ellington A.G., Frankland J.A., Frazer A., French L., Garner P.,  
 Graiham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 Levasialho M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,  
 Marsh V.L., Martin S.L., McConnachie L.J., McLeay K., McMurray A.A.,  
 Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 Rice C.D., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,  
 Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
 Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
 Rogers J.;  
 RL "The DNA sequence and comparative analysis of human chromosome 20.";  
 Nature 414:865-871(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=Kidney, and Placenta;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 Fahey J., Helton E., Kettman M., Maman A., Rodriguez S., Sanchez A.,  
 Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,  
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences".  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
 RN [5]  
 RP SEQUENCE OF 83-903 FROM N.A. (ISOFORM 2), AND INTERACTION WITH DRPLA.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=98313405; PubMed=9647693;  
 RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,  
 Kaminsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.;  
 RT "Atrophin-1, the DRPLA gene product, interacts with two families of WW  
 RT domain-containing proteins".  
 RL Mol. Cell. Neurosci. 11:149-160(1998).  
 RN [6]  
 RP SEQUENCE OF 463-470; 503-510; 514-526; 644-665 AND 875-881,  
 RP INTERACTION WITH LMP2A, AND MUTAGENESIS OF CYS-871.  
 RC TISSUE=B-cell;  
 RX MEDLINE=20501262; PubMed=11046148;  
 RA Winberg G., Matekova I., Chen F., Plant P., Rotin D., Gish G.,  
 Ingham R., Ernberg I., Pawson T.;  
 RT "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3  
 RT protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases".  
 RL Mol. Cell. Biol. 20:8526-8535(2000).  
 RN [7]  
 RP INTERACTION WITH CBLC, AND PHOSPHORYLATION.  
 RX MEDLINE=22323223; PubMed=12226085;  
 RA Courbard J.-R., Fiore F., Adelaide J., Borg J.P., Birnbaum D.,  
 Ollendorff V.;  
 RT "Interaction between two ubiquitin-protein isopeptide ligases of  
 RT different classes, CBLC and AIP4/ITCH".  
 RL J. Biol. Chem. 277:45267-45275(2002).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and



RN SEQUENCE FROM N.A., AND INTERACTION WITH WBP1; WBP2; SCNN1A; SCNN1B  
 RP AND SCNN1G.  
 RC TISSUE=Bone marrow, and Brain;  
 RX MEDLINE=97313427; PubMed=9169421;  
 RA Pirozzi G., McConnell S.J., Uveges A.J., Carter J.M., Sparks A.B.,  
 RA Kay B.K., Fowlkes D.M.;  
 RT "Identification of novel human WW domain-containing proteins by  
 RT cloning of ligand targets";  
 RL J. Biol. Chem. 272:14611-14616(1997).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Placenta;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickens M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RN INTERACTION WITH DRPLA, AND TISSUE SPECIFICITY.  
 RX MEDLINE=98313405; PubMed=9647693;  
 RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan X., Kushi J.,  
 RA Kaminsky Z., Kleiderlein J.J. Jr., Sharp A.H., Ross C.A.;  
 RT "Atrophin-1, the DRPLA gene product, interacts with two families of  
 RT WW domain-containing proteins";  
 RL Mol. Cell. Neurosci. 11:149-160(1998).  
 RN [4]  
 RN INTERACTION WITH SCNN1A, SCNN1B AND SCNN1G.  
 RX MEDLINE=22157111; PubMed=12167593;  
 RA McDonald F.J., Western A.H., McNeil J.D., Thomas B.C., Olson D.R.,  
 RA Snyder P.M.;  
 RT "Ubiquitin-protein ligase WBP2 binds to and downregulates the  
 RT epithelial Na(+) channel";  
 RL Am. J. Physiol. 283:F431-F436(2002).  
 RN [5]  
 RN INTERACTION WITH ADENOVIRUS TYPE 2 PIII.  
 RX MEDLINE=22338357; PubMed=12450395;  
 RA Galinier R., Gout E., Lortat-Jacob H., Wood J., Chroboczek J.;  
 RT "Adenovirus protein involved in virus internalization recruits  
 RT ubiquitin-protein ligases";  
 RL Biochemistry 41:14295-14305(2002).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts  
 CC ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a  
 CC thioester and then directly transfers the ubiquitin to targeted  
 CC substrates (By similarity).  
 CC -!- SUBUNIT: Binds SCNN1A, SCNN1B, SCNN1G, WBP1, DRPLA and  
 CC adenovirus type 2 PIII.  
 CC -!- TISSUE SPECIFICITY: Detected in heart, throughout the brain,  
 CC placenta, lung, liver, muscle, kidney and pancreas.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 4 WW domains.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase  
 CC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U96114; AAC51325.1; -;  
 DR EMBL; BC000108; AAH00108.1; -;  
 DR EMBL; BC013645; AAH13645.1; -;  
 DR HSSP; Q13526; IPIN.  
 DR MIM; 602308; -;  
 DR GO; GO:0000151; C:ubiquitin ligase complex; TAS.  
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; TAS.  
 DR GO; GO:0006464; P:protein modification; TAS.  
 DR InterPro; IPR000008; C2\_CaLB.  
 DR InterPro; IPR008973; C2\_CaLB.  
 DR InterPro; IPR000569; HECT\_domain.  
 DR InterPro; IPR002349; WW.  
 DR Pfam; PF00168; C2; 1.  
 DR Pfam; PF00632; HECT; 1.  
 DR Pfam; PF00397; WW; 4.  
 DR PRINTS; PR00403; WWDOMAIN.  
 DR SMART; SM00239; C2; 1.  
 DR SMART; SM00119; HECTC; 1.  
 DR SMART; SM00456; WW; 4.  
 DR PROSITE; PS00499; C2\_DOMAIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50004; C2\_DOMAIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50237; HECT; 1.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; 4.  
 DR PROSITE; PS50020; WW\_DOMAIN\_2; 4.  
 DR Ubl conjugation pathway; Ligase; Repeat.  
 KW Ubl conjugation pathway; Ligase; Repeat.  
 FT DOMAIN 20 100 C2 DOMAIN.  
 FT DOMAIN 300 333 WW 1.  
 FT DOMAIN 330 363 WW 2.  
 FT DOMAIN 405 437 WW 3.  
 FT DOMAIN 444 477 WW 4.  
 FT DOMAIN 536 870 HECT.  
 FT BINDING 838 838 UBIQUITIN (BY SIMILARITY).  
 FT BINDING 136 136 E -> K (IN REF. 1).  
 FT CONFLICT 394 395 SS -> FW (IN REF. 1).  
 SQ SEQUENCE 870 AA; FCCD75CBA61F2204 CRC64;  
 Query Match 32.2%; Score 1298.5; DB 1; Length 870;  
 Best Local Similarity 42.0%; Pred. No. 8.7e-85;  
 Matches 271; Conservative 100; Mismatches 217; Indels 57; Gaps 9;  
 QY 122 PNNDVTVRGQIVVLSQSRDRI GTGQGVDCSR LFDNDLPDGWEERTASGRIOYLNHITR 181  
 DB 265 PNNTTSLPAPATPAEGEPSSGTQQLPAAQAPDLPAGWEORLPNGRVVYVDHTK 324  
 QY 182 TTOWERPTRPASYSFGRPLSCFVDENTPI SGTNATCGQSSDPPLAERRVRSQRHNY 241  
 DB 325 TTTWERPLPGWEKRTDPRGRFYVDHNT-----RTTTWQRTAEYVRNVEQWQSRNQ 378  
 QY 242 MSRTLLH-----TPPD-----LPEGYEQR TCGQGVFLHTGTGVTWHDPRV 284  
 DB 379 LQAMQHFSORFLYQSSASTDHDPLGPPGWEKE-QDNGRVVYVNNHNTTRITQWEDPT 437  
 QY 285 PRDLSNINCELGPLPPGWEIRNTATGRVYFVDHNNRTQTFTDRLSANLHLVLNQNL 344  
 DB 438 QGMQIE-----PALPGWEMKYTSGVAYFVDHNTTRITTFKDPFGFE----- 480  
 QY 345 KDQQQQVSLCPDDTECLTPYKEDLVOKLILRQELSQQQPQAGHCRIVSRREIFE 404  
 DB 481 -----SGTKQSGPGAYDRSFRWKYQFR-FLCHSNALPSHKIVSRQTLFE 526  
 QY 405 ESYRQVMKMRPKDLWKRLMKFRGEGDLYGGVAREWLYLLSHEMLNPNYVGLFOYGRDDI 464  
 DB 527 DSFQIINWMPYDLRRLRLYIMRGEGLDYGGIAREWFFLLSHEVLNPNVCLFEYAGKN 586

QY 465 YTLQINPDSAVNPEHLSYFHFVGRINGMAVFHGHYIDGGFTLPFYKQLLGKSTLDDMEL 524  
 DB 587 YCQINPASSINDHLTYFRIQGFAMADYHGKFDITGFTLPFYKRLMKKPTLKDLES 646  
 QY 525 VDPDLHNSLWILENDITGV-LDHTFCVEHAYGEIIQHELKFNKSGSIPVNEENKEYVR 583  
 DB 647 IDPEFYNSIVWIKENNLEOCCLELYFIQDMELIGKVTTHLKEGGSIRVTSENKEYIM 706  
 QY 584 LYVWNRFLRGIEAFLAQGFNEV:POHLLKTFDEKELELLICGLGKIDVDKNTL 643  
 DB 707 LLTDWRFTRGVEBQTKAFLGDFNEVAPLEWLRVDFDEKELELMCGQOEIDMSDWQKSTIY 766  
 QY 644 KHTCPSNIVKFWKAVFEFFDERRERLLOFVTVGSSRVLPQGFALQGAAGRPRLFTIHOI 703  
 DB 767 RHYTKNSKQIOWVQVQVWKNEDNEKRILOFVTVGTCRLPVGGFAELIGSNGKQKFCIDKV 826  
 QY 704 DACTNNLPKATCFNRIDIPPYESYKLYEKLKLTALTEBCTCGFAVE 748  
 DB 827 GKET-WLPSRHTCFNRLDLPYKSYBQLREKLYAIEETEGFQGE 870

RESULT 13  
 WP1\_HUMAN  
 ID WP1\_HUMAN STANDARD; PRT; 922 AA.  
 AC Q9H0M0; O00307; Q96BP4;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Nedd4-like ubiquitin-protein ligase WBP1 (EC 6.3.2.-) (WW domain-  
 DE containing protein 1) (Atropin-1 interacting protein 5) (AIP5).  
 GN WBP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Testis;  
 RA MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 RA Ansoerg W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
 RA Mewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,  
 RA Wambutt R., Korn B., Klein M., Foutska A.;  
 RA "Towards a catalog of human genes and proteins: sequencing and  
 RA analysis of 500 novel complete protein coding human cDNAs.";  
 RL Genome Res. 11:422-435(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5 AND 6), AND TISSUE  
 RP SPECIFICITY.  
 RC TISSUE=Breast cancer;  
 RX MSDLINE=21638010; PubMed=11779188;  
 RX Flasz M., Gorman P., Roylance R., Canfield A.E., Baron M.;  
 RT "Alternative splicing determines the domain structure of WBP1, a  
 RT Nedd4 family protein.";  
 RL Biochem. Biophys. Res. Commun. 290:431-437(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Lung, and Testis;  
 RX MSDLINE=22388257; PubMed=12477932;  
 RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavani T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton A., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butcherfield Y.S.N., Krzywinski W.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.N., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 191-870 FROM N.A., AND INTERACTION WITH WBP1; WBP2;  
 RP SCNN1A; SCNN1B AND SCNN1G.  
 RC TISSUE=Bone marrow, and Brain;  
 RX MEDLINE=97313427; PubMed=9169421;  
 RA Pirozzi G., McConnell S.J., Uveges A.J., Carter J.M., Sparks A.B.,  
 RA Kay B.K., Fowlkes D.M.;  
 RT "Identification of novel human WW domain-containing proteins by  
 RT cloning of ligand targets.";  
 RL J. Biol. Chem. 272:14611-14616(1997).  
 RN [5]  
 RP INTERACTION WITH DRPLA, AND TISSUE SPECIFICITY.  
 RX MEDLINE=98313405; PubMed=9647693;  
 RA Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,  
 RA Kaminsky Z., Kleiderlein J.J. Jr., Sharp A.H., Ross C.A.;  
 RT "Atrophin-1, the DRPLA gene product, interacts with two families of  
 RT WW domain-containing proteins.";  
 RL Mol. Cell. Neurosci. 11:149-160(1998).  
 RN [6]  
 RP INTERACTION WITH PIIL.  
 RX MEDLINE=22338357; PubMed=12450395;  
 RA Galanter R., Gout E., Lortat-Jacob H., Wood J., Chroboczek J.;  
 RT "Adenovirus protein involved in virus internalization recruits  
 RT ubiquitin-protein ligases.";  
 RL Biochemistry 41:14299-14305(2002).  
 RN [7]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 546-917, MUTAGENESIS OF  
 RP GLU-614; HIS-621; ASP-675; GLU-798; MET-804; ARG-845; GLN-848  
 RP AND ARG-855, AND FUNCTION.  
 RX MEDLINE=22423789; PubMed=12535537;  
 RA Verdecia M.A., Joazeiro C.A.P., Wells N.J., Ferrer J.-L., Bowman M.E.,  
 RA Hunter T., Noel J.P.;  
 RT "Conformational flexibility underlies ubiquitin ligation mediated by  
 RT the WBP1 HECT domain E3 ligase.";  
 RL Mol. Cell 11:249-259(2003).  
 CC -!- FUNCTION: E3 ubiquitin-protein ligase which accepts ubiquitin from  
 CC an E2 ubiquitin-conjugating enzyme in the form of a thioester and  
 CC then directly transfers the ubiquitin to targeted substrates.  
 CC -!- PATHWAY: Ubiquitin conjugation; third step.  
 CC -!- SUBUNIT: Binds KLF2 (By similarity). Binds SCNN1A, SCNN1B, SCNN1G,  
 CC WBP1, WBP2, DRPLA and adenovirus type 2 PIIL.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=6;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1; Synonyms=A;  
 CC IsoId=Q9H0M0-1; Sequence=Displayed;  
 CC Name=2; Synonyms=B;  
 CC IsoId=Q9H0M0-2; Sequence=VSP\_007601, VSP\_007603;  
 CC Name=3; Synonyms=C;  
 CC IsoId=Q9H0M0-3; Sequence=VSP\_007602;  
 CC Name=4; Synonyms=D;  
 CC IsoId=Q9H0M0-4; Sequence=Not described;  
 CC Name=5; Synonyms=E;  
 CC IsoId=Q9H0M0-5; Sequence=Not described;  
 CC Name=6; Synonyms=F;  
 CC IsoId=Q9H0M0-6; Sequence=VSP\_007600;  
 CC -!- TISSUE SPECIFICITY: Detected in heart, placenta, pancreas, kidney,  
 CC liver, skeletal muscle, bone marrow, fetal brain, and at much  
 CC lower levels in adult brain and lung. Isoforms 1 and 5 predominate  
 CC in all tissues tested, except in testis and bone marrow, where  
 CC isoform 5 is expressed at much higher levels than isoform 1.  
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-  
 CC thiolester formation.  
 CC -!- SIMILARITY: Contains 1 C2 domain.  
 CC -!- SIMILARITY: Contains 4 WW domains.  
 CC -!- SIMILARITY: Contains 1 HECT-type E3 ubiquitin-protein ligase







```

Db      856  GFAELMGNGPKQKCIKVGKDT-WLPRSHCTCFNRDLDPYKSVQKEXLLFAIBETEG 91L
      745  FAVE 748
      |
Db      915  FGQE 918

```

Query Match	31.7%	Score 1281	DB 1	Length 918
Best Local Similarity	44.9%	Pred. No. 1.7e-034		
Matches 271	Conservative 93	Mismatches 194	Indels 46	Gaps 12
QY	159	LPDGWEERTASGRIQYLNHITRTTOWER--PTRPASEYSSPCRLSCFVDENTPISGTN	216	
Db	347	LPSGWEQRKDPHGRITVYVDHNTTRITTTWERQPLPGWERRVDRGRVYVDHNTTRITTTQ	406	
QY	217	GATC-----CGSSDPPLAERVRVSQRHNYM-SRTHLTPPD-----LPEGVEQRTTOG	265	
Db	407	RETMSVRNFFQWOSQRNQCGAMQQNQBYLYSASMLAAENDPYGPPPGWEKRVSDTD	466	
QY	266	QYVFLLHTQTGVSTWHDPVRPRLDSNINCEBGLPLPGWEIRNTATGRVYFVDHNNRTTQF	325	
Db	467	RVYFVNHTKTIQWEDPRT-QGLEF--EE--PLPGWEIRYTRGVRVYVDHNTTRITTF	520	
QY	326	TPRUSANLHLVNRQNLKDDQQQQVVSICPDDTECLTVPRYKRLVOKLILRQELSQ	385	
Db	521	KDPR-----NGKSVTKGGPQIA-----YERFRWKLAAHFY-LCQ	555	
QY	386	QQPQAGHCKRIEVSREITFEESYROWVMKMRKDLWKLMIKFRGEGLDGGVAREWLXLL	445	
Db	556	SNALPSHWKINVSRQTLFEDSFQOIMALKPYDLRRLLYVIFRGEGLDYGGLAREWFFLL	615	
QY	446	SHEMLNPPYGFPQYSRDDIYTLQINPDASVNPBLSYFHFVGRINGMAVFGHHYIDGGFT	505	
Db	616	SHEVLNPMYCLFVAGKKNYCLQINPASTINPDLSYFCFICRGITAMALFHGKFI DTGFS	675	
QY	506	LPFYKOLLGKSTITDDMLVDPDILHNSLVMLINDITGV-LDHTFCVEHNAVEIIOHEL	564	
Db	676	LPFYKRLMSKLLTIKLESITETPYSNLIIWRDNTEECGLEMYFSVDMELIGKVTSHDL	735	
QY	565	KPGKSIIPVNEKNKYYRLVYNNRFLRGIEAQLAQGFNEVTPQHLLKTFDSEKLEL	624	
Db	736	KLGGSNILVTBENKDEYTLGTMETNRFSGVQEQTKAFLDQFNVEWPQLQWLFDFEKELEV	795	
QY	625	ITCGLGKIDVNDVKNTLELKHCTPDSNI VKWFKAVEFFDEERARLLQFTVGSRRVPLQ	684	
Db	796	MLCGHEQVDLADQRNTVYRHYTNSKQIINFVQCFVKETDNEVYMELLOQFTVGTCELPLG	855	
QY	685	GFKALQGAAGPRLFTIHQIDACTNNLKAHTCFNRIDIPPYSESYEKLYBKLLTAIEETCG	744	

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:47:37 ; Search time 95.0891 Seconds  
(without alignments)  
2526.134 Million cell updates/sec

Title: US-10-009-945-4

Perfect score: 4038

Sequence: 1 MSNPGRRNGPVKRLTVLC.....EKLYEKLTAETTCGFAVE 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCr\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCrUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4031	99.8	748	US-10-021-660-81	Sequence 81, Appl
2	4027	99.7	804	US-10-220-120-407	Sequence 407, App
3	3939.5	97.6	735	US-10-313-955-2	Sequence 2, Appli
4	3001	74.3	722	US-10-097-534-14	Sequence 14, Appl
5	2775	68.7	514	US-09-764-875-819	Sequence 819, App
6	1564	38.7	766	US-10-313-955-4	Sequence 4, Appli
7	1530.5	37.9	832	US-10-032-585-7296	Sequence 7296, Ap
8	1453.5	36.0	869	US-10-128-714-8162	Sequence 8162, Ap
9	1370.5	33.9	921	US-10-205-823-279	Sequence 279, App
10	1355.5	33.6	725	US-10-185-050-126	Sequence 126, App
11	1354.5	33.5	854	US-10-205-823-277	Sequence 277, App
12	1354.5	33.5	927	US-10-097-534-15	Sequence 15, Appl
13	1353.5	33.5	995	US-10-097-534-9	Sequence 9, Appli
14	1353.5	33.5	995	US-10-205-823-275	Sequence 275, App
15	1353	33.5	834	US-10-313-955-6	Sequence 6, Appli

16	1340.5	33.2	759	14	US-10-128-714-3162	Sequence 3162, Ap
17	1338.5	33.1	854	14	US-10-287-218-3	Sequence 3, Appli
18	1338.5	33.1	854	16	US-10-474-291-3	Sequence 3, Appli
19	1301	32.2	739	12	US-10-182-936A-89	Sequence 89, Appl
20	1301	32.2	739	14	US-10-097-534-10	Sequence 10, Appl
21	1301	32.2	739	15	US-10-374-979-89	Sequence 89, Appl
22	1301	32.2	752	10	US-09-919-039-235	Sequence 235, App
23	1298.5	32.2	870	14	US-10-097-534-12	Sequence 12, Appl
24	1298.5	32.2	906	14	US-10-185-050-48	Sequence 48, Appl
25	1228	30.4	898	12	US-10-188-186-114	Sequence 114, App
26	1209	29.9	474	10	US-09-774-639-371	Sequence 371, App
27	1209	29.9	474	10	US-09-969-730-249	Sequence 249, App
28	1209	29.9	474	16	US-10-621-363-249	Sequence 249, App
29	1134.5	28.1	683	14	US-10-185-050-46	Sequence 46, Appl
30	1134.5	28.1	684	14	US-10-097-534-11	Sequence 11, Appl
31	1050	26.0	1562	14	US-10-097-534-13	Sequence 13, Appl
32	1019.5	25.2	380	14	US-10-307-956-32	Sequence 32, Appl
33	1008.5	25.0	375	14	US-10-307-956-31	Sequence 31, Appl
34	869	21.5	733	14	US-10-097-534-16	Sequence 16, Appl
35	835.5	20.7	1094	14	US-10-043-487-300	Sequence 300, App
36	835.5	20.7	1488	14	US-10-043-487-285	Sequence 285, App
37	835.5	20.7	2011	16	US-10-408-765A-767	Sequence 767, App
38	832.5	20.6	1104	12	US-10-424-599-146002	Sequence 146002, App
39	828.5	20.5	1843	12	US-10-424-599-146004	Sequence 146004, App
40	828.5	20.5	3647	16	US-10-437-963-119793	Sequence 119793, App
41	791	19.6	157	12	US-10-211-462-107	Sequence 107, App
42	775.5	19.2	277	9	US-09-925-300-1628	Sequence 1628, Ap
43	771	19.1	973	16	US-10-437-963-178922	Sequence 178922, App
44	733	18.2	358	14	US-10-268-036-5	Sequence 5, Appli
45	733	18.2	358	16	US-10-391-364-33	Sequence 33, Appl

#### ALIGNMENTS

#### RESULT 1

US-10-021-660-81  
; Sequence 81, Application US/10021660  
; Publication No. US20030152926A1  
; GENERAL INFORMATION:  
; APPLICANT: Murray, Richard  
; APPLICANT: Glynn, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: EOS Biotechnology, Inc.  
; TITLE OF INVENTION: No. US20030152926A1 Methods of Diagnosis of Angiogenesis,  
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis  
; TITLE OF INVENTION: Modulators  
; FILE REFERENCE: 018501-000710US  
; CURRENT APPLICATION NUMBER: US/10/021,660  
; CURRENT FILING DATE: 2001-12-06  
; PRIOR APPLICATION NUMBER: US/09/784,356  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: US 09/637,977  
; PRIOR FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 81  
; LENGTH: 748  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-021-660-81

Query Match 99.8%; Score 4031; DB 14; Length 748;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 747; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MSNPGRRNGPVKRLTVLCACNLVKKDFRLLPDPFAKVVDGSGGCHSDTIVKNTLDPK 60  
Db 1 MSNPGRRNGPVKRLTVLCACNLVKKDFRLLPDPFAKVVDGSGGCHSDTIVKNTLDPK 60  
Qy 61 WNCHYLDYIGKSDSVTISVWNHKKHKOGAGFLGCVRLLSNAINRLKDTGYORLDCKL 120  
Db 61 WNCHYLDYIGKSDSVTISVWNHKKHKOGAGFLGCVRLLSNAINRLKDTGYORLDCKL 120

```
121 GPNDNTRVGOIIVVLSQSRDRIGTGGVQVDCSRLFNDLPDGEWEERTASGRIOYLNHIT 180
121 GPNDNTRVGOIIVVLSQSRDRIGTGGVQVDCSRLFNDLPDGEWEERTASGRIOYLNHIT 180
181 RTTQWERPTRPASEYSPGRPLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN 240
181 RTTQWERPTRPASEYSPGRPLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN 240
241 YMSRTHLHTPPDLPEGYEQRTTQGGVYFLHTGTGVSTWHDPRVPRDLNINCEELGFLP 300
241 YMSRTHLHTPPDLPEGYEQRTTQGGVYFLHTGTGVSTWHDPRVPRDLNINCEELGFLP 300
301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQLKDDQOQQVVSCLPDDT 360
301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQLKDDQOQQVVSCLPDDT 360
361 ECLTVPRYKRDVLQVKILRQELSSQOQOAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
361 ECLTVPRYKRDVLQVKILRQELSSQOQOAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
421 RLMTKFRGEGLDYGVARWELVLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
421 RLMTKFRGEGLDYGVARWELVLLSHEMLNPPYGLFOYSRDDIYTLQINPDSAVNPEHL 480
481 SYHFVGRIMGMAVFGHYIDGGFTLPFYKOLLGKSIITLDDMELVDPDLHNSLVILEND 540
481 SYHFVGRIMGMAVFGHYIDGGFTLPFYKOLLGKSIITLDDMELVDPDLHNSLVILEND 540
541 ITGVLDHTFCVEHNAYGEIIOHELKPKNGKSI PVNEENKEYVRLVYNNWRFLRGIEAQFLA 600
541 ITGVLDHTFCVEHNAYGEIIOHELKPKNGKSI PVNEENKEYVRLVYNNWRFLRGIEAQFLA 600
601 LQKGFNEVPOHLKLTDEKELELIICGLGKIDVNDWKVNTRLKCHCTPDSNIVKFWKAV 660
601 LQKGFNEVPOHLKLTDEKELELIICGLGKIDVNDWKVNTRLKCHCTPDSNIVKFWKAV 660
661 EFFDEERRARLQVFGSSRVPLOGFKALQOAGPRFTTHQIDACTNNLPKHAHTCFNRI 720
661 EFFDEERRARLQVFGSSRVPLOGFKALQOAGPRFTTHQIDACTNNLPKHAHTCFNRI 720
721 DIPVESYEKLYEKLTAIESTCGFAVE 748
721 DIPVESYEKLYEKLTAIESTCGFAVE 748

RESULT 2
US-10-220-120-407
; Sequence 407, Application US/10220120
; Publication No. US20040048253A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: PANZER, Scott R.
; APPLICANT: SPIRO, Peter A.
; APPLICANT: BANVILLE, Steven C.
; APPLICANT: SHAH, Purvi
; APPLICANT: CHALUP, Michael S.
; APPLICANT: CHANG, Simon C.
; APPLICANT: CHEN, Alice
; APPLICANT: D'SA, Steven A.
; APPLICANT: AMSHEY, Stefan
; APPLICANT: DAHL, Christopher R.
; APPLICANT: DAM, Tam C.
; APPLICANT: DANIELS, Susan E.
; APPLICANT: DUFOUR, Gerard E.
; APPLICANT: FLORES, Vincent
; APPLICANT: FONG, Willy T.
; APPLICANT: GREENAWALT, Lila B.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: JONES, Anissa L.
; APPLICANT: LIU, Tommy F.
; APPLICANT: ROSEBERRY, Ann M.
; APPLICANT: ROSEN, Bruce H.

; APPLICANT: RUSSO, Frank D.
; APPLICANT: STOCKREHER, Theresa K.
; APPLICANT: DAFPO, Abel
; APPLICANT: WRIGHT, Rachel J.
; APPLICANT: YAP, Pierre E.
; APPLICANT: YU, Jimmy Y.
; APPLICANT: BRADLEY, Diana L.
; APPLICANT: BRATCHEY, Shawn R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: COHEN, Howard J.
; APPLICANT: HODGSON, David M.
; APPLICANT: LINCOLN, Stephen E.
; APPLICANT: JACKSON, Stuart
; TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PT-1113 FCT
; CURRENT APPLICATION NUMBER: US/10/220,120
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 60/184,777; 60/184,777; 60/184,777; 60/184,698; 60/184,770; 60/184,774;
; 60/184,693; 60/184,771; 60/184,813; 60/184,773; 60/184,776;
; 60/184,769; 60/184,768; 60/184,837; 60/184,697; 60/184,841;
; 60/184,772; 60/185,213; 60/185,216; 60/204,863; 60/205,221;
; 60/204,815; 60/203,785; 60/204,821; 60/204,908; 60/204,226;
; 60/204,525; 60/205,285; 60/205,232; 60/205,323; 60/205,287;
; 60/205,324; 60/205,286
; PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24; 2000-02-24;
; 2000-05-17; 2000-05-12; 2000-05-16; 2000-03-16; 2000-05-15;
; 2000-05-16; 2000-05-17; 2000-05-16; 2000-05-17; 2000-05-17;
; 2000-05-17; 2000-05-17
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PERL Program
; SEQ ID NO 407
; LENGTH: 804
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040048253A1 LG:132147.3.crf3:2000FEB18
US-10-220-120-407

Query Match 99.7%; Score 4027; DB 12; Length 804;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSNPGRRNGPVKRLTLVLCANLVKKDFRLPDPFAKVVVDGSGQCHSTDTVXNTLDPK 60
Db 57 MSNPGRRNGPVKRLTLVLCANLVKKDFRLPDPFAKVVVDGSGQCHSTDTVXNTLDPK 116
QY 61 MNQHYDLVIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL 120
Db 117 MNQHYDLVIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQLDLCKL 176
QY 121 GPNDNTRVGOIIVVLSQSRDRIGTGGVQVDCSRLFNDLPDGEWEERTASGRIOYLNHIT 180
Db 177 GPNDNTRVGOIIVVLSQSRDRIGTGGVQVDCSRLFNDLPDGEWEERTASGRIOYLNHIT 236
QY 181 RTTQWERPTRPASEYSPGRPLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN 240
Db 237 RTTQWERPTRPASEYSPGRPLSCFVDENTPISGTNGATCGOSSDPRLAERRVRSQRHN 296
QY 241 YMSRTHLHTPPDLPEGYEQRTTQGGVYFLHTGTGVSTWHDPRVPRDLNINCEELGFLP 300
Db 297 YMSRTHLHTPPDLPEGYEQRTTQGGVYFLHTGTGVSTWHDPRVPRDLNINCEELGFLP 356
QY 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQLKDDQOQQVVSCLPDDT 360
Db 357 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQLKDDQOQQVVSCLPDDT 416
QY 361 ECLTVPRYKRDVLQVKILRQELSSQOQOAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420
Db 417 ECLTVPRYKRDVLQVKILRQELSSQOQOAGHCRIEVSREEIFEESYRQVMKRPKDLWK 476
```

QY 421 RLMIKRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480  
DB 477 RLMIKRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 536  
QY 481 SYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILEND 540  
DB 537 SYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILEND 596  
QY 541 ITGVLDHTFCVHNAYGEIIQHELKPKNGKSIIPVNEENKKEYVRLYVNNRFLRGIEBAQFLA 600  
DB 597 ITGVLDHTFCVHNAYGEIIQHELKPKNGKSIIPVNEENKKEYVRLYVNNRFLRGIEBAQFLA 656  
QY 601 LQKGFNEVPIQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660  
DB 657 LQKGFNEVPIQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 716  
QY 661 EFFDEERRARLLQFVTGSSRVLPQGFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 720  
DB 717 EFFDEERRARLLQFVTGSSRVLPQGFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 776  
QY 721 DIPPYESYEKLYEKLTAIEETCGFAVE 748  
DB 777 DIPPYESYEKLYEKLTAIEETCGFAVE 804

RESULT 3  
US-10-313-955-2  
; Sequence 2, Application US/10313955  
; Publication No. US20030199036A1  
; GENERAL INFORMATION:  
; APPLICANT: Beach, David H.  
; Nefsky, Bradley  
; Caligiuri, Maureen  
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/10/313,955  
; FILING DATE: 05-Dec-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/392,163  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 08/539,205  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: CSV-005.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-313-955-2  
Query Match 97.6%; Score 3939.5; DB 14; Length 735;

Best Local Similarity 98.0%; Pred. No. 0;  
Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;  
QY 1 MSNFGRRNGPVKLRILVLCALNVLKXDFRLLPDPFAKVVVDSGGQCHSTDTVKNTLDPK 60  
DB 1 MSNFGRRNGPVKLRILVLCALNVLKXDFRLLPDPFAKVVVDSGGQCHSTDTVKNTLDPK 47  
QY 61 WNQHYDYIGKSDSVTISVMNHKKIHKKGAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 120  
DB 48 WNQHYDYIGKSDSVTISVMNHKKIHKKGAGFLGCVRLLSNAINRLKDTGYQRLDLCKL 107  
QY 121 GPNDNDTVRGQIVVLSQSRDRIGTGGQVDCSRLFDNDLPGHEERTASGRTOYLNHIT 180  
DB 108 GPNDNDTVRGQIVVLSQSRDRIGTGGQVDCSRLFDNDLPGHEERTASGRTOYLNHIT 167  
QY 181 RTTOWERPTTPASYSYSPGRLSCFVDENTPIISGTNGATCGQSDPPLAERRVRSQRHN 240  
DB 168 RTTOWERPTTPASYSYSPGRLSCFVDENTPIISGTNGATCGQSDPPLAERRVRSQRHN 227  
QY 241 YMSRTHLHTPPDLPEGYEQRITQGGQVYFHTQTGVSTWHDPRVPRDLSNINCEELGPLP 300  
DB 228 YMSRTHLHTPPDLPEGYEQRITQGGQVYFHTQTGVSTWHDPRVPRDLSNINCEELGPLP 287  
QY 301 PGWEIRNTATGRVYFVDHNNRTTQFDPRLSANLHLVNEQNLKDOQQQVVSCLCPDDT 360  
DB 288 PGWEIRNTATGRVYFVDHNNRTTQFDPRLSANLHLVNEQNLKDOQQQVVSCLCPDDT 347  
QY 361 ECLTVPRYKDLVQKILQELSQOQOQAGHCRIEVSREEIFEESYRQVMKRPKDLWK 420  
DB 348 ECLTVPRYKDLVQKILQELSQOQOQAGHCRIEVSREEIFEESYRQVMKRPKDLWK 407  
QY 421 RLMIKRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 480  
DB 408 RLMIKRGEGLDYGGVAREWLYLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAVNPEHL 467  
QY 481 SYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILEND 540  
DB 468 SYFHFVGRIMGMAVFHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILEND 527  
QY 541 ITGVLDHTFCVHNAYGEIIQHELKPKNGKSIIPVNEENKKEYVRLYVNNRFLRGIEBAQFLA 600  
DB 528 ITGVLDHTFCVHNAYGEIIQHELKPKNGKSIIPVNEENKKEYVRLYVNNRFLRGIEBAQFLA 587  
QY 601 LQKGFNEVPIQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 660  
DB 588 LQKGFNEVPIQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWFKAV 647  
QY 661 EFFDEERRARLLQFVTGSSRVLPQGFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 720  
DB 648 EFFDEERRARLLQFVTGSSRVLPQGFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 707  
QY 721 DIPPYESYEKLYEKLTAIEETCGFAVE 748  
DB 708 DIPPYESYEKLYEKLTAIEETCGFAVE 735  
RESULT 4  
US-10-097-534-14  
; Sequence 14, Application US/10097534  
; Publication No. US20030049607A1  
; GENERAL INFORMATION:  
; APPLICANT: GREENER, TSVIKA  
; APPLICANT: MOSKOWITZ, HAIM  
; APPLICANT: REISS, YUVAL  
; APPLICANT: ALROY, IRIS  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL  
; FILE REFERENCE: PLV-001.01  
; CURRENT APPLICATION NUMBER: US/10/097,534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 60/275,224  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/308,958

APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: RJ202  
CURRENT APPLICATION NUMBER: US/09/764,875  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1249  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 819  
LENGTH: 514  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-875-819

Query Match 68.7%; Score 2775; DB 11; Length 514;  
Best Local Similarity 99.8%; Pred. No. 2.3e-248;  
Matches 513; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 235 SORHNTMSRTHLHTPPDLPEGYEORTTQOGVYFLHTGTGVSHTWHPVPRDLNINCE 294  
Db 1 SRRHNTMSRTHLHTPPDLPEGYEORTTQOGVYFLHTGTGVSHTWHPVPRDLNINCE 60  
QY 295 ELGPLPGWEIRNTATGRVYFVDHNNETTQTPRLSANLHLVLRNQNLKDDQQQQVVS 354  
Db 61 ELGPLPGWEIRNTATGRVYFVDHNNETTQTPRLSANLHLVLRNQNLKDDQQQQVVS 120  
QY 355 LCPDTECLTVPRYKRDVQKILROELSSQQOQAGHCRIEVSREIPEESYRQVMKMR 414  
Db 121 LCPDTECLTVPRYKRDVQKILROELSSQQOQAGHCRIEVSREIPEESYRQVMKMR 180  
QY 415 PKDLWKRLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 474  
Db 181 PKDLWKRLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 240  
QY 475 VNPEHLSYFHFVGRIMGMAVPHGHYIDGGFTLPYKQLLGSITLDDMELVDDPLHNSLV 534  
Db 241 VNPEHLSYFHFVGRIMGMAVPHGHYIDGGFTLPYKQLLGSITLDDMELVDDPLHNSLV 300  
QY 535 WILENDITGVLDHFTFCVEHNAAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFMRGI 594  
Db 301 WILENDITGVLDHFTFCVEHNAAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFMRGI 360  
QY 595 BAOFLALQKGFNEVTPQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVK 654  
Db 361 BAOFLALQKGFNEVTPQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVK 420  
QY 655 WFKAVEFFDEERRARLLQFVTGSRVPLQGFKALQGAAGPRFLTIIHQAIDACTNNLPK 714  
Db 421 WFKAVEFFDEERRARLLQFVTGSRVPLQGFKALQGAAGPRFLTIIHQAIDACTNNLPK 480  
QY 715 TCFNRIDIPPPESYEKLEKLLTAIEETCGFAVE 748  
Db 481 TCFNRIDIPPPESYEKLEKLLTAIEETCGFAVE 514

RESULT 6  
US-10-313-955-4  
Sequence 4, Application US/10313955  
Publication No. US20030199036A1  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
Nefsky, Bradely  
Caligiuri, Maureen  
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:

PRIOR FILING DATE: 2001-07-31  
PRIOR APPLICATION NUMBER: 60/340,170  
PRIOR FILING DATE: 2001-12-07  
NUMBER OF SEQ ID NOS: 71  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 14  
LENGTH: 722  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-097-534-14

Query Match 74.3%; Score 3001; DB 14; Length 722;  
Best Local Similarity 74.6%; Pred. No. 4.1e-269;  
Matches 565; Conservative 63; Mismatches 71; Indels 58; Gaps 8;

QY 12 VKLRLTVLCAKLVKDDFFRLPDPFAKVVVDGSGQCHSTDTVKNLTDPKWNQHYLDYVGK 71  
Db 4 IKRLTVLCAKLVKDDFFRLPDPFAKVVVDGSGQCHSTDTVKNLTDPKWNQHYLDYVGK 63  
QY 72 SDSVTSVWNHKKIHKKGAGFGCVRLLSNAINRLKDTGYQRLDCKLGPNDNDTVRGQ 131  
Db 64 TDSITTSVWNHKKIHKKGAGFGCVRLLSNAINRLKDTGYQRLDCKLNPSTDAVRGQ 123  
QY 132 IVVSLOQRDRIGTGGQVVDGSRFLPDNDLPDQWEERETASGRIOYLNHITRTTQWERPTRP 191  
Db 124 IVVSLOQRDRIGTGGVVDGSRGLENE-----GTVY----- 154  
QY 192 ASEYSSGRLSPCHVDENTPISGTNGATCG-----QSSPRLAERRVRSQRHNTYM 242  
Db 155 --EDSGRPLSCFMEPSAPYTDSTGAAGGNGCRFVESPQORLOAQRNPDVVRGSL 212  
QY 243 ----SRTHLTPDPLPGVEORTTQOGVYFLHTGTGVSHTWHPVPRDLNINCEELGP 298  
Db 213 QTPQNRHGHQSPLPGVEORTTQOGVYFLHTGTGVSHTWHPVPRDLNINCEDELGP 272  
QY 299 LPPGWEIRNTATGRVYFVDHNNETTQTPRLSANLHLVLRNQNLKDDQQQQVVS-----S 354  
Db 273 LPPGWEIRNTATGRVYFVDHNNETTQTPRLSANLHLVLRNQNLKDDQQQQVVS-----S 328  
QY 355 LCPDTECLTVPRYKRDVQKILROELSSQQOQAGHCRIEVSREIPEESYRQVMKMR 414  
Db 329 L---EDBELPQVRVERDLQKILROELSSQQOQAGHCRIEVSREIPEESYRQVMKMR 385  
QY 415 PKDLWKRLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 474  
Db 386 PKDLWKRLMKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTLQINPDSA 445  
QY 475 VNPEHLSYFHFVGRIMGMAVPHGHYIDGGFTLPYKQLLGSITLDDMELVDDPLHNSLV 534  
Db 446 INPDHLSYFHFVGRIMGMAVPHGHYIDGGFTLPYKQLLGSITLDDMELVDDPLHNSLV 505  
QY 535 WILENDITGVLDHFTFCVEHNAAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFMRGI 594  
Db 506 WILENDITGVLDHFTFCVEHNAAYGIIQHELKPNKSIIPVNEENKKEYVRLYVNWRFMRGI 565  
QY 595 BAOFLALQKGFNEVTPQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVK 654  
Db 566 BAOFLALQKGFNEVTPQHLKTFDEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVK 625  
QY 655 WFKAVEFFDEERRARLLQFVTGSRVPLQGFKALQGAAGPRFLTIIHQAIDACTNNLP 714  
Db 626 WFKAVEFFDEERRARLLQFVTGSRVPLQGFKALQGAAGPRFLTIIHQAIDACTNNLP 685  
QY 712 KAHTCFNRIDIPPPESYEKLEKLLTAIEETCGFAVE 748  
Db 686 KAHTCFNRIDIPPPESYEKLEKLLTAIEETCGFAVE 722

RESULT 5  
US-09-764-875-819  
Sequence 819, Application US/09764875  
Publication No. US20040018969A1  
GENERAL INFORMATION:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/313.955
; FILING DATE: 05-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/392.163
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/539,205
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-313-955-4

Query Match      38.7%; Score 1564; DB 14; Length 766;
Best Local Similarity 41.1%; Pred. No. 1.4e-135;
Matches 338; Conservative 121; Mismatches 232; Indels 132; Gaps 18;

QY      1 MSNPGRRNGPVKRLTLVCAKNLVKDFRLDPFAKVVVDSGGCHGCHTDVTKNLTDPK 60
DB      1 MSNSAQR-----RIRVTVAADGLYKRDVFRFPDPFAVLTVDGE-QTHTTTAKTLNPEY 55
QY      61 WNOHYDLYICKDSVTISVWNHKKIKKQAGFLGCVRLLSNAINELKDTGYQLR--DLG 118
DB      56 WNETFEVNTDNTIAIQVFDQKFF-KKKQGQFLGVINRVGDVLDAIGGDEMLTRDK 114
QY      119 KLGPNNDTVRGQIVVSLQ-----
DB      115 K--SNENTVVGKIIINLTAAQTILQVPSSAASGARTQTSITNDPQSSKSSVSNNPA 172
QY      138 -----SRDRTGTGGQVDCSRLEPN-----DLPDGWEERTASGRIOYLNH:TRTQ 184
DB      173 SSRAGSPTDNDAPASPASBPRTSSFDQGRPPPGWERDNLGRYYVYDHNTRSTT 232
QY      185 WERPT-----RPASYSPGRLSCFVDENTPIGCTNGATCGSSDPRLAERRVRSQRH 238
DB      233 WIRPNLLSVAGAAAEHLSSAS--SANVTEGVQPSSSNAA-----RRTASVLT--- 279
QY      239 RNYMSRTHLTPDLPPEGVEQRTQQGVYFLHTQTGVSTWHDPRVPRDLNIN----- 292
DB      280 ----SNATTAGSGELPPGWEQRTYEGRFYVDHNTRITTTWDPRRQYRSYGGFNNA 335
QY      293 -----CEEGLPLPPGWEIRNTATGRVYVDHNTRITQTDPRLSANHLVLNQNQLKDX 347
DB      336 IQQQPVSLGQLPSPGWEMLTNTARYVVDHNKTKTTWDDPLPSSL-----DQ 384
QY      348 QQQQVSLCPDTECLTPRYKRDILVQKLIURQLSQ--QPQAGHCRIEVSRREEIFE 405
DB      385 -----NVFQYKRDFFRKLIYF---LSQPALHPLPGQCHIKVRNHFIED 425
QY      406 SYRQVMQMRPKDLWKRLMKRFGESGLDYGGVAREWLVLLSHEMLNPNYGLFOYSRDDIY 465
DB      426 SYAEIMRQSATDLKRLMKFDGEGDLGYGGSUREYFLLSHEMFNPFYCLFEYSSVDNY 485
QY      466 TLQINPDSAVNPEHLSYFHFVGRINGMAVFHGHIYDGGFTLFPYKQLLGKSIITLDDMELV 525
DB      486 TLQINPHSGINPEHNLNFKYGRVIGLAIFHRFVDFAFFVSVFYQWILQKVTQLQDESM 545
QY      526 DDDLNSLWILENDITGVLDHTFCVEHNAIGEIIQHELKPKNGKSIIPVNEENKKEYVELY 585

```

## RESULT 7

```

US-10-032-585-7296
; Sequence 7296, Application US/10032585
; Publication No. US20030180953A1
; GENERAL INFORMATION:
; APPLICANT: Terry, Roemer D.
; APPLICANT: Bo, Jiang
; APPLICANT: Charles, Boone
; APPLICANT: Howard, Bussey
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug Target Discovery
; FILE REFERENCE: 10182-005-999
; CURRENT APPLICATION NUMBER: US/10/032.585
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7296
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Candida albicans
US-10-032-585-7296

```

```

Query Match      37.9%; Score 1530.5; DB 14; Length 832;
Best Local Similarity 38.4%; Pred. No. 2.1e-132;
Matches 338; Conservative 120; Mismatches 228; Indels 195; Gaps 21;

```

```

QY      6 RRRNGPVKRLTLVCAKNLVKDFRLDPFAKVVVDSGGCHGCHTDVTKNLTDPKYNQHY 65
DB      9 RSNNTTINVKVVAASLYKRDVFRFPDPFAVLTVDGS-QTKTITAKKTLNPNWNETF 67
QY      66 DLYIGKSDSV--TISVWNHKKIKKQAGFLGCVR-----LLSNAINELKDTGYQLDLCK 119
DB      68 N-FOAKEDSILVIQVPQKXF-KKQDQFLGVINRVIGDVIDLSNASEETIR--DLKK 123
QY      120 LGPNDNDTVRGQIVVSLQSRDRIGTGGQVVD-----
DB      124 --SNENLAVSGKIIVI--SHNRNSNGGVTTATTGTGASSNNIATITSGVNNLRIGSA 180
QY      151 -----
DB      181 TTTANSTAQASSDATVANGSGTSSLPPIGQHPETAATPGGAAGAAASQYSSFD 240
QY      158 ---DLPDGWEERTASGRIOYLNH:TRTTOWERPTRPASE----- 194
DB      241 QYGLPLPGWERRTDFNFGRTYYVDHNSRTTTWQRPALHQSETERGQQRQSQSEAEARRHRG 300
QY      195 YSPGPR---PLSCFVDENTPIGCTNGATCGSS-----DPRLAERRVRSQRHN 240
DB      301 RTLPGEQSVSL-----PTSGNSITSGNVTVNAGSANTPVNPAAVSMAAGATTS 352
QY      241 YMSRTHLTPDLPPEGVEQRTQQGVYFLHTQTGVSTWHDPRVPRDLNIN----- 292
DB      353 GLG-----ELPSGWEQRTTEGRPYVDHNTRITTTWDPRRQYVIRTFGNTIIQQQ 404
QY      293 -CEEGLPLPGWEIRNTATGRVYVDHNTRITQTDPRLSANHLVLNQNQLKQDQOOQQ 351
DB      405 PVSQGLPLPSGWEMRLNTARVYVDHNKTKTTWDDPLPSSL-----DQ----- 449

```

QY 352 VVSLCPDDTECLTVPRYKRDVLQKILRQELSSQOPO-----AGHCRIEVSREBIFRESY 407  
DB 450 -----NVPOYKDDFRKVIYFR-----SQPALRILPGQCHIKVRDRHIFEDSY 492  
QY 408 ROVMKMRPKDLWKRLMIKFRGEGLDYGGVAREWYLLSHEMLNPPYGLFOYSRDDIYTL 467  
DB 493 QETMRQTPEDLKKRLMIKFRGEGLDYGGVAREWYLLSHEMLNPPYGLFOYSRDDIYTL 552  
QY 468 QINPDGAVANPEHLSYFHFVGRIMGMAVPHGHVIDGGFTLPYKOLLGKSIITLDDMELVDP 527  
DB 553 QINPNGINPEHLSYFHFVGRIMGMAVPHGHVIDGGFTLPYKOLLGKSIITLDDMELVDP 612  
QY 528 DLNLSVLWILENDITGVLDHTFCVHNAYGEITIOHELKPNKSGIPVNEENKKEYVRYLYN 587  
DB 613 EFRSLKWLINDITGILDLTFSABEESFGEIVEVDLPGGDRDIEVTENKHEYVELITE 672  
QY 588 WRFLRGIEAQLAQGFNEVPOHLKTFDEKELELIIICGLIKTDVNDWKNVRLKCHT 647  
DB 673 WRISKEVEQFKAFIDGFMELIPQELVNVDFDERELLELLIGLAEIDCEDWKKHTDYRGYQ 732  
QY 648 POSNIVKFWKAVFPEERRARLLQFVTGSRVPLQGFALQGAAGPRFTIHOIDACT 707  
DB 733 ENDQVQFWKCNINEMDSQKARLLQFVTGSRVPLQGFALQGAAGPRFTIHOIDACT 791  
QY 708 NNLPKAHTCFNRIDIPYSEYKLEKLTALTEETCGFAVE 748  
DB 792 NQLPKSHTCFNRVDPFPPYDYSLKQKLTALVEETVGFQGE 832

## RESULT 8

US-10-128-714-8162  
; Sequence 8162, Application US/10128714  
; Publication No. US20030119013A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Bo  
; APPLICANT: Hu, Wenqi  
; APPLICANT: Tishkoff, Daniel  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Broshkin, Alexey M  
; APPLICANT: Lemieux, Sebastien M  
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and  
; TITLE OF INVENTION: Methods of Use  
; FILE REFERENCE: 10182-018-999  
; CURRENT APPLICATION NUMBER: US/10128,714  
; CURRENT FILING DATE: 2002-04-23  
; PRIOR APPLICATION NUMBER: US 60/285,697  
; PRIOR FILING DATE: 2001-04-23  
; PRIOR APPLICATION NUMBER: US 60/287,066  
; PRIOR FILING DATE: 2001-04-27  
; PRIOR APPLICATION NUMBER: US 60/295,890  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/303,899  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/316,362  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 8603  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 8162  
; LENGTH: 869  
; TYPE: PRT  
; ORGANISM: Aspergillus fumigatus  
US-10-128-714-8162

Query Match 36.0%; Score 1453.5; DB 14; Length 869;  
Best Local Similarity 36.2%; Pred. No. 3.2e-125;  
Matches 332; Conservative 114; Mismatches 229; Indels 243; Gaps 21;  
QY 11 PVKRL-TVLCANLVKKDDFFR-----LPDPKAVVDDGSOCH 48  
DB 15 PASSRLCRVIAADGLYKRDVFRKSVILAILLVRLTSGAFAGFPDPFAVATVGGE-QTH 73  
QY 49 STDVTAKNTLDPKWNHVDYLDKSDVTSIVNWHKIKHKGAGFLGCVRL----- 100

DB 74 TTSVIKTLNPNVNMFMWRVNESSILAIQIFQOKKF-KKKQOGLGVINVRIGDVIDLQ 132  
QY 101 -----SNAINPLK-----DTGYORLDLCKLGPNDNDTVRGQIVWSL----- 136  
DB 133 MGGDGESLPIRHSQDVSRLLTFRLLLETEMLETRDLKK--SNDNLVHGKLIILNLSLST 190  
QY 137 -----QSRDRIGTGQVWDCS----- 152  
DB 191 PNTNQANGLHRSHVQSSSTSSGLVQVAPSSHPAASGTA PVDPSASNPSPNLPQVPSTTR 250  
QY 153 -----RLPDND-----LPQWEERTASRIQIYLNHIIRTTQ 184  
DB 251 PSTAAPASAAAVNSHSGSRITNLSFESQGLPAGWERREDNLGRTTYVDHNTTTT 310  
QY 185 WERPTPASEYSSPGRPISCFVDENTPISGTNGATCGQSSDPRLAERRVRSQRHRYM-- 242  
DB 311 W---TRPSSNYNEHAQ-----RSQREANWQLERRAHOQRMPLP 344  
QY 243 -----SRTHLHTTP-----DLPEGYQRTTQGOVY 268  
DB 345 EDRTGANSNPPESSQQAHTPPAGGSANAVSMATGATTAGTGLPQGWQRTTPEGRPY 404  
QY 269 FLHTGTGVSTWHDPRVPRDL-----SNIN-----CEELGPLPQGWIRNTATGRVY 314  
DB 405 FVDHNTTRTTTVDPRQYIRMYGNANGTNTTQQQPVSQLGLPGLSGWEMRLTNTARVY 464  
QY 315 FVDHNTTRTTQTDRLSANLHLVLRQNLKXQOQOQVSVLCPDDTECLTVPRYKRDVQ 374  
DB 465 FVDHNTKTTTWDPRPLPSSL-----DQ-----GVPOYKRDFFR 497  
QY 375 KLIKRLQELSSQOQ-----OAGHCRIEVSREBIFRESYROVMKMRPKDLWKRLMIKFERGE 430  
DB 498 KLIYFR-----SQPALRIMSQCCHVKYRNNIFEDSYAEIMRQASDCLKKRLMIKFDGED 552  
QY 431 GLDYGAVAREWYLLSHEMLNPPYGLFOYSRDDIYTLQINPDGAVANPEHLSYFHFVGRIM 490  
DB 553 GLDYGSLSGEFFFLLSHEMLNPPYGLFOYSRDDIYTLQINPDGAVANPEHLSYFHFVGRV 612  
QY 491 GMAVPHGHVIDGGFTLPYKOLLGKSIITLDDMELVDPDLNLSVLWILENDITGVLDHTFC 550  
DB 613 GLAIFHRFLOSFFTGAFYKWLKRYSLQDMGVEDDLHRLNTWTWDDNDIEGVELITS 672  
QY 551 VEHNAYGEITIOHELKPNKSGIPVNEENKKEYVRYLYNWRFLRGIEAQLAQGFNEVIP 610  
DB 673 VDDEKFGERRTIDILKPGGDRDIPVTNENKASYRLVTEWKIVKVEEQFNAFMSGFNELIP 732  
QY 611 QHLKTFDEKELELIIICGLIKTDVNDWKNVRLKCHTCTPDSNIVKFWKAVFPEERRAR 670  
DB 733 ADLVNVFDERELLELLIGIADIIVDDAKKHTDYRGYQESDEVIONFWKIVRSWDAEQKSR 792  
QY 671 LLQFVTGSRVPLQGFALQGAAGPRFTIHOIDACTNNLPKATCFNRIDIPYSEYK 730  
DB 793 LLQFVTGSRVPLQGFALQGAAGPRFTIHOIDACTNNLPKATCFNRIDIPYSEYK 730  
QY 731 LYEKLLTAIETTCGFAVE 748  
DB 852 LEHKNMSIAVENTLGFQGE 869

## RESULT 9

US-10-205-823-279  
; Sequence 279, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Monsey, Angela M.  
; APPLICANT: Glatt, Karen

APPLICANT: Zhao, Xumei  
APPLICANT: Anderson, Dustin  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
THERAPY OF PROSTATE CANCER  
FILE REFERENCE: MRI-044  
CURRENT APPLICATION NUMBER: US/10/205,823  
CURRENT FILING DATE: 2002-07-25  
PRIORITY FILING DATE: 2002-07-25  
PRIORITY FILING DATE: 2001-07-25  
PRIORITY FILING DATE: 2001-08-22  
PRIORITY FILING DATE: 2001-09-25  
PRIORITY FILING DATE: 2001-12-12  
PRIORITY FILING DATE: 2001-12-12  
PRIORITY FILING DATE: 2002-03-05  
NUMBER OF SEQ ID NOS: 455  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 279  
LENGTH: 911  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-205-823-279

Query Match 33.9%; Score 1370.5; DB 14; Length 911;  
Best Local Similarity 36.7%; Pred. No. 1.8e-117;  
Matches 324; Conservative 119; Mismatches 234; Indels 207; Gaps 28;

QY 24 LVKDDFFELP---DPEAKV---VVDGSGOCH---STDVTKNTLPKNOHYDLYIGKSD-S 74  
DB 69 LPTDELFPGICDPYKLSLYADENRELALVQTKTKLTPKKNWEEFYFVNPNSNR 128  
QY 75 VTISVNNHKKIKKQAGFLGCVR-----LLSNAINRLKDTGY 112  
DB 129 LLFEVFDENLRD---FLQVDVPLSHLPTEDPTMERYTEFKDILLRPSHKSRVKG 185  
QY 113 ORLDLCKLGNNDTVRGQIVSVLSQSDRIGTGGVQVDCSGLPDND-----L 159  
DB 186 LRLKMYAPKNG-----GQBEESDDQDDMEHGEWVD-----SNDASQHQEELPPPL 235  
QY 160 PDGWEERTASGRQYLNHTRTTQWERPT-----RPA 192  
DB 236 PPGWEKVDNLGRYYVNNHNRITQWRPSLMDVSSDNNIRQINQEAHRRFRSRH1 295  
QY 193 SEYSSP-----GRPLSCFVDENTISGTN-----GATCGOSSDPR-LAE---RE 232  
DB 296 SEDLEPESGDDVPEPWET-ISEVNIAGSLGLALPPPPASGERTSPQELSELSRR 354  
QY 233 VR-----SORHNYMSRTHL-----HTPPD-----252  
DB 355 LQITPDSNGEQFSSLIQREPSRLRSCSVTDVAEQQHLPFGAKDGPVRRRAVKDTLSNPQ 414  
QY 253 -----LPEGYEORTTQCGQVYELHTQTGVSTWHDPRVP-----RD 287  
DB 415 SPQSPYNSPKQHKVTSQSPFPQWEMRIAPNGPFFIDHNTKTTWEDPRLKFPYHMS 474  
QY 288 LSNINCEELGELPGWEIRNTATGRVYFVDHNRITQDTPRLSANHLVNLNQKLDQ 347  
DB 475 KTSLNPNDLGFLPGWEERIHLDGRTFYDHSKITQWEDPRL-----QN-----519  
QY 348 QQQQVSVLCPDDTECLTVPR--YKEDLVOKLILRQELSQQQQAQAGHREIVSREIPEE 405  
DB 520 -----PAITGPAPVYREPKQKYDYFRKKLKKPADIPNRFENKLRHNNIFEE 566  
QY 406 SYRQVMKMRPKDLWK-RIMIKFRGEGLDYGGVAREWLYLLSHMLNPNYYGLFQYSRDDI 464  
DB 567 SYRRIMSVRPDVLKARLWIEFESEKGLDYGGVAREWFFLLSKEMFNYYGLFYSATDN 626  
QY 465 YTLQINPDSAV-NPEHLSYFFHVRINGMAVFGHYIDGGFTLPFYKQLIGKSLTLDME 523  
DB 627 YTLQINPNPNSGLCNEDHLSYFTFIRGVAGLAVFGHKLDDGFFIRPFYKQMLGKQITLDME 686

QY 524 LVDPDLHNSLVWILENDITGVLDHTTCVHNAYGELIQHLEKPKNGKSIIPVNEENKKEYVR 583  
DB 687 SVDSYNSLKWILENDPT-ELDSLMFCIDEENFGQTYQVYDLKPKNGSEIMVTNENKREYID 745  
QY 584 LYVNRRLRGIEAQFLALOKGFNEVIPHQLLTKTFDEKELELIICGLGKIDVNDKVNTRL 643  
DB 746 LVTCMRVNRVQKMNAFLEGTTELLPIDLIKIFDENELLEMLCGLGDVVDVNDWRQHSY 805  
QY 644 K-HCTPDSNIVKFWKAVEFFDEERRARLQVTCSSRVPLQGFKALOGAGPLFTTH 701  
DB 806 KNGYC-PNHFVIOFWKAVALLMDAERLLOFVOTGTSRVPMNGFAELYGSNGPOLFTIE 864  
QY 702 QIDACTNNLKAHTCFNRIDIPPEYSEYKLYEKLTAIETCGF 745  
DB 865 QWGS-PEKLPRHTCNRLDLPYETFEFLREKLLMAVENAOGF 907

## RESULT 10

US-10-185-050-126  
Sequence 126, Application US/10185050  
Publication No. US2003007577A1  
GENERAL INFORMATION:  
APPLICANT: Pirozzi, Gregorio  
Kay, Brian K.  
Fowlkes, Dana M.

TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME

NUMBER OF SEQUENCES: 233  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/185,050  
FILING DATE: 28-Jun-2002  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/826,516  
FILING DATE: 03-Apr-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: MISROCK, S. LESLIE

REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-208-999

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090  
TELEFAX: (212) 896-8864/9741  
TELEX: 66141 PENNIE

## INFORMATION FOR SEQ ID NO: 126:

## SEQUENCE CHARACTERISTICS:

LENGTH: 725 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: unknown

MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 126:

US-10-185-050-126

Query Match 33.6%; Score 1355.5; DB 14; Length 725;  
Best Local Similarity 43.7%; Pred. No. 3.1e-116;  
Matches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;

QY 137 QSRDLCTGQGVDCSLFON---DLPGDWEERTASGRLOYNLHTRTTQWERPRPAS 193  
DB 112 RARSSTVTGGEETPPSVAYVHTTFLGSPGWEERKAKGRTYVNNHNRITTTWTRPINQUA 171

```
194 EYSSPG-----RPLSC---FVDENTPISTNGATCGOSSDPRLAERVRVS 235
172 EDGASGATNNHLLIEPQIRRRSLSSPTVTLAPLEGAKDSPVRAVKDTLSNPQSPQ 231
236 QHRNYSMTLHTPPDLPEGVEORTTQCGOVYFLHTGTGVSTWHDPRVP-----RDLN 290
232 PPSYNSPKQHKVTSQSLPPGWEIRAPNGRFFFDHNTKTITWEDPRLKFPVHMRKTS 291
291 INCEELGPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLNRQNLKQDQOQ 350
292 LNPDLGPPGWEIRHLDGRTFYIDHNSKITQWEDPRL-----QN----- 333
351 QVVSCLPDDTECLTVPR--YKEDLVOKLKLRLQELSSQOQAGHCRIEVSREIFESYR 408
334 -----PAITGPAVPSREFKQKYDYFRKLLKPADIPNRFEMKLRNNIFESYR 512
409 QVMKMRPKDLWK-RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTL 467
384 RIMSVKRPDVLKARLWIEFESEKGLDYGGVAREWFFLLSKEMENFYGLFEYSATDNYTL 443
468 QINPDSAV-NPEHLSYFHFVGRINGMAVPHGHYIDGGFTLPRYKOLLGKSTLDDMELVD 526
444 QINPNSGLCNEDHLSYFTFGRVAGLAVPHGKLDGFFIRFFYKMLGKQITLNDMESVD 503
527 PDLHNSLWILENDITGVLDHTFCVEHNAYGEIIQHELKPNKGKSIIPVNEENKKEYVRLYV 586
504 SEYNSLKWILENDPT-ELDLMFCDISENFGQTYQVDLKPNGSEIMVTNENKREYIDLVI 562
587 NWRFLRGIEAQLALQKGFNEVIPHOLLKTDFEKELELIICGLKIDVNDKVNTRLK-- 644
563 QMRFVNRVQKQMNAPLEGFTTELLPDLIKIFDENELEMLCGLDGVDVNDWROHSYKNG 622
645 HCTPDSNIVKMFKAVFEFFDEERRARLLQFVTGSSRVPLQGFKALQGAAGPRLFTIHQID 704
623 YC-PNHPVIOQWFKAVLLMDAERIRLLQFVTGTSRVPMNGFAELYGSGNGPQLFTIEQWG 681
705 ACTNPLPKAHTCFNRIDIPPYESYKLYEKLITAEETCGF 745
682 S-PEKLPKAHTCFNRIDLPPEYTFEDLREKLLMAVENAQGF 721
```

## RESULT 11

```
US-10-205-823-277
; Sequence 277, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbacheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Womsey, Angela M.
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
```

```
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 277
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-205-823-277

Query Match 33.5%; Score 1354.5; DB 14; Length 854;
Best Local Similarity 43.7%; Pred. No. 5e-116;
Matches 280; Conservative 98; Mismatches 200; Indels 63; Gaps 13;

QY 137 QSRDRIGTCGGVDCSRLFDN---DLPDGEERTASGRIOYLNHITTTQWERPREPAS 193
Db 241 RARSSTVTGSGEPTSVAYVHTTTPGLPSGWEERKDAKRTYYVNNHNRITTTWTRPIMQLA 300
QY 194 EYSSPG-----RPLSC---FVDENTPISTNGATCGOSSDPRLAERVRVS 235
Db 301 EDGASGATNNHLLIEPQIRRRSLSSPTVTLAPLEGAKDSPVRAVKDTLSNPQSPQ 360
QY 236 QHRNYSMTLHTPPDLPEGVEORTTQCGOVYFLHTGTGVSTWHDPRVP-----RDLN 290
Db 361 PPSYNSPKQHKVTSQSLPPGWEIRAPNGRFFFDHNTKTITWEDPRLKFPVHMRKTS 420
QY 291 INCEELGPPGWEIRNTATGRVYFVDHNNRTTQTDPRLSANLHLVLNRQNLKQDQOQ 350
Db 421 LNPDLGPPGWEIRHLDGRTFYIDHNSKITQWEDPRL-----QN----- 462
QY 351 QVVSCLPDDTECLTVPR--YKEDLVOKLKLRLQELSSQOQAGHCRIEVSREIFESYR 408
Db 463 -----PAITGPAVPSREFKQKYDYFRKLLKPADIPNRFEMKLRNNIFESYR 512
QY 409 QVMKMRPKDLWK-RLMIKFRGEGLDYGGVAREWLYLLSHEMLNPPYGLFOYSRDDIYTL 467
Db 513 RIMSVKRPDVLKARLWIEFESEKGLDYGGVAREWFFLLSKEMENFYGLFEYSATDNYTL 572
QY 468 QINPDSAV-NPEHLSYFHFVGRINGMAVPHGHYIDGGFTLPRYKOLLGKSTLDDMELVD 526
Db 573 QINPNSGLCNEDHLSYFTFGRVAGLAVPHGKLDGFFIRFFYKMLGKQITLNDMESVD 632
QY 527 PDLHNSLWILENDITGVLDHTFCVEHNAYGEIIQHELKPNKGKSIIPVNEENKKEYVRLYV 586
Db 633 SEYNSLKWILENDPT-ELDLMFCDISENFGQTYQVDLKPNGSEIMVTNENKREYIDLVI 691
QY 587 NWRFLRGIEAQLALQKGFNEVIPHOLLKTDFEKELELIICGLKIDVNDKVNTRLK-- 644
Db 692 QMRFVNRVQKQMNAPLEGFTTELLPDLIKIFDENELEMLCGLDGVDVNDWROHSYKNG 751
QY 645 HCTPDSNIVKMFKAVFEFFDEERRARLLQFVTGSSRVPLQGFKALQGAAGPRLFTIHQID 704
Db 752 YC-PNHPVIOQWFKAVLLMDAERIRLLQFVTGTSRVPMNGFAELYGSGNGPQLFTIEQWG 810
QY 705 ACTNPLPKAHTCFNRIDIPPYESYKLYEKLITAEETCGF 745
Db 811 S-PEKLPKAHTCFNRIDLPPEYTFEDLREKLLMAVENAQGF 850
```

## RESULT 12

```
US-10-097-534-15
; Sequence 15, Application US/10097534
; Publication No. US20030049607A1
; GENERAL INFORMATION:
; APPLICANT: GREENER, TSVIKA
; APPLICANT: MOSKOWITZ, HAIM
; APPLICANT: REISS, YUVAL
; APPLICANT: ALROY, IRIS
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
; FILE REFERENCE: PLV-001.01
; CURRENT APPLICATION NUMBER: US/10/097,534
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/275,224
```



QY 552 EHNAYGEIIQHELKPNKSGIPVNEENKKEVYRLVYNNWFLRGIEAQLALQKGFNEVIPO 611  
DB 798 DEENFGQTYQVLLKPNKSGSEIMVTNENKREYDLVQWRFVNRVQKQNAFLGFTLELPI 857  
QY 612 HLKTFDEKELELEIIICGLGKIDVNDKWNTRLK--HCTPDSNIVKWFVKAVERFDEBERRA 669  
DB 858 DLKIFDENLELELLMCGLDVNDWRQHSIYKNGYC-PNHPVQWFWKAVLLDAEKRI 916  
QY 670 RLLOFVTGSSRPVLOQFKALQAGAPRLFTIHQIDACTNNLPKATCFNRIDIPPEYSYE 729  
DB 917 RLLOFVTGSSRPVPMNGFAELYSNGPOLFTIEQWGS-PEKLPRAHCFNRDLPPYETFE 975  
QY 730 KLYEKLTAIBETQGF 745  
DB 976 DLREKLLMAVENAQQF 991  
RESULT 14  
US-10-205-823-275  
; Sequence 275, Application US/10205823  
; Publication No. US20030108963A1  
; GENERAL INFORMATION:  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Endege, Wilson O.  
; APPLICANT: Gannavarapu, Manjula  
; APPLICANT: Gorbacheva, Bella  
; APPLICANT: Hoersch, Sebastian  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Wonsay, Angela M.  
; APPLICANT: Glatt, Karen  
; APPLICANT: Zhao, Xumei  
; APPLICANT: Anderson, Dustin  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
; FILE REFERENCE: MRI-044  
; CURRENT APPLICATION NUMBER: US/10/205, 823  
; CURRENT FILING DATE: 2002-07-25  
; PRIOR APPLICATION NUMBER: 60/307, 982  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: 60/314, 356  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: 60/325, 020  
; PRIOR FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: 60/341, 746  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/362, 158  
; PRIOR FILING DATE: 2002-03-05  
; NUMBER OF SEQ ID NOS: 455  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 275  
; LENGTH: 995  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-205-823-275  
Query Match 33.5%; Score 1353.5; DB 14; Length 995;  
Best Local Similarity 44.6%; Pred. NO. 7.9e-116;  
Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;  
QY 159 LPDGEERTASGRIOYLNHTRTOWERPTRPASEYSSPO-----RPLS 203  
DB 407 LPSGEERKADGRYYVYNNHRTTTPFIMQALDAGSASATNSNNHLEIQRPRPS 466  
QY 204 C---FVDENTPISGNGATCGOSSDPRLAERVRVRSQRHNNVMSSTHHTTPDLPGEYEQR 260  
DB 467 LSSPTVTL-SAPLEGAKDSFVRANVADTLNSPQSPSPYNSPKQKHVTSQFLPPGWEMR 526  
QY 261 TTQOQGVFLHQTGVSTWHPDVP-----RDLSNINCEELGPLPPQWEIRNTATGRVVF 315  
DB 527 IAPNGRPFDHNTKTITWEDPRKFPVHMSKTSLSNPDLGPLPPGWERIHLDGTFY 586

QY 316 VDHNRRTQTTDPRLSANLHLVLRNQNLKDDQQQQQVSLCPDDTECLTVPR--YKRDLV 373  
DB 587 IDHNSKITQWSDPRL-----QN-----PAITGPAVPYSREPK 618  
QY 374 OKKLILROELSQOQOQACHCRIEVSREEIFEESYRQWKMPPKDLWK-RLMIKFRGEEGL 432  
DB 619 QKYDYFRKLLKPADIPNRPFKLHRNNIFEESYRIMS VKRPDVLKARLWIEFESEKGL 678  
QY 433 DYGVGAREWLVLLSHEMLNPPYGLFQYSRDDIYTLQINPDSAV-NPEHLSYFHFVCRING 491  
DB 679 DYGVGAREWVFLLSKEMFNPPYGLFEYSATDNVTLQINPNSGJCNEHLSYFTFICRVAG 738  
QY 492 NAVPHGHYIDGGFTLPFYKQLLGKSIITLDDMELVDPDLHNSLVWILENDITGVLDHTFCV 551  
DB 739 LAVFHGKLLDGGFFIRPFYKMLGKQITLNDWESVDSEYNSLWILENDPT-ELDLMFCI 797  
QY 552 EHNAYGEIIQHELKPNKSGIPVNEENKKEVYRLVYNNWFLRGIEAQLALQKGFNEVIPO 611  
DB 798 DEENFGQTYQVLLKPNKSGSEIMVTNENKREYDLVQWRFVNRVQKQNAFLGFTLELPI 857  
QY 612 HLKTFDEKELELEIIICGLGKIDVNDKWNTRLK--HCTPDSNIVKWFVKAVERFDEBERRA 669  
DB 858 DLKIFDENLELELLMCGLDVNDWRQHSIYKNGYC-PNHPVQWFWKAVLLDAEKRI 916  
QY 670 RLLOFVTGSSRPVLOQFKALQAGAPRLFTIHQIDACTNNLPKATCFNRIDIPPEYSYE 729  
DB 917 RLLOFVTGSSRPVPMNGFAELYSNGPOLFTIEQWGS-PEKLPRAHCFNRDLPPYETFE 975  
QY 730 KLYEKLTAIBETQGF 745  
DB 976 DLREKLLMAVENAQQF 991  
RESULT 15  
US-10-313-955-6  
; Sequence 6, Application US/10313955  
; Publication No. US20030199036A1  
; GENERAL INFORMATION:  
; APPLICANT: Beach, David H.  
; APPLICANT: Caligiuri, Maureen  
; APPLICANT: Nefsky, Bradley  
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/313,955  
; FILING DATE: 05-Dec-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/09/392,163  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: US 08/539,205  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: CSV-005.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 834 amino acids

```
;
;      TYPE: amino acid
;      TOPOLOGY: linear
;      MOLECULE TYPE: protein
;      SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-313-955-6

Query Match      33.5%; Score 1353; DB 14; Length 834;
Best Local Similarity 44.6%; Pred. No. 6.6e-116;
Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;

QY 159 LPDQWERRTASGRIQVNLHITTTOWERPTRPASEYSPG-----RPLS 203
Db 246 LPSGWEERKDAKRTYYVNNHNTTTTTPIMQAEADGASGATNSNNHLLIEPQIRRRPS 305
QY 204 C--FVDENTPIGTCGATCGQSSDPRLAERRVRSQRHRYNMSRTHLHTPPDLPEGYEQR 260
Db 306 LSSPTVTLSPLEGAKDSPVRRAVKXTLSNPQSPQSPYNSPKPKVKVTSQFLPPGWNKR 365
QY 261 TTQCGQVYFHTGTGTVSTWHDPRVP-----RDLNSINCEBELGFLPPGWEIRNTATGRVYF 315
Db 366 IAPNGRPFPIIDHNKTITTTWEDPRLKFPVHWRSKTSLNPNDLGFLPPGWEIRHLDGRIFY 425
QY 316 VDHNNRTTQFTDPRLSANLHLVLRNQNLKDOQQQVSLCPDDTECLTVPR--YKRDIV 373
Db 426 IDNSKITQWEDRL-----QN-----PAITGPAPVYSREPK 457
QY 374 QKLKILRQELSQOQPOAGHCRIEVSREIEFESYQVWKMVRPKDLWK-RLMIKFRGEGL 432
Db 458 QKYDYFRKKLKPADIPNPFEMKLRHNNIFESYRIRMSVKRPDLVKARLWIEFSEKGL 517
QY 433 DYGGVAREWLYLLSHEMLNPYGLFCYSRDDIYTIQINPDSAV-NPEHLSYFHFVGRIMG 491
Db 518 DYGGVAREWFFLLSKEMFNPYGLFEYSATDNYTLQINPNSGLCNEDHLSYFTFGRVAG 577
QY 492 MAVFHGHIYDGGFTLPFYKQLLCKSITLDDMELVDPDLHNSLVWILENDITGVLDHTFCV 551
Db 578 LAVFHGKLDGFFIRPFYKWLKGKQITLNDMESVDSEYNSLKWILENDPT-ELDLMFCI 636
QY 552 EHNAYGEITQHELKPKNGKSI PVNEENKKEYVLYVNRFLRGIEAQFLALQKGFNEVIPQ 611
Db 637 DEENFGQTYQVDLKPNGSEIMVTNENKREYIDLVIQWRFVNRVQKQNAFLGFTTELLPI 696
QY 612 HLKTFDEKELELIICGLGKIDVNDWKVNTLK--HCTPDSNIVKWKVKEVFFDEERRA 669
Db 697 DLIKIFDENELLCGLGDVDVNDWRQHSIYKNGYC-PNHEVIOQWFKAVLLMDAEKRI 755
QY 670 RLLOFVTGSSRVPLOGKALQAGAPRLFTIHQIDACTNNLPKAHTCFNRIDIPPEYSYE 729
Db 756 RLLOFVTGTSRVPMNGFAELYGNGPQLFTTBQWGS-PEKLPRAHTCFNRDLDPPEYTFE 814
QY 730 KLYEKLLTAIETCGF 745
Db 815 DLREKLLMAVENAQGF 830
```

Search completed: September 21, 2004, 08:04:28  
Job time : 100.589 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 07:40:45 ; Search time 23.8994 Seconds  
(without alignments)  
1615.783 Million cell updates/sec

Title: US-10-009-945-4

Perfect score: 4038

Sequence: 1 MSNPGRRNGPVKRLTVLC.....EKLYEKLTAIBETCGFAVE 748

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A COMB pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCFUS COMB pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1 pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3939.5	97.6	735	3	US-08-539-205A-2
2	3939.5	97.6	735	4	US-08-332-163A-2
3	1564	38.7	766	3	US-08-539-205A-4
4	1564	38.7	766	4	US-08-332-163A-4
5	1354.5	33.5	927	3	US-08-895-601-6
6	1353	33.5	834	3	US-08-539-205A-6
7	1353	33.5	834	4	US-08-332-163A-6
8	1336	33.1	854	2	US-08-070-060-4
9	1336	33.1	854	3	US-08-357-746-4
10	1329	32.9	852	2	US-08-070-060-3
11	1329	32.9	852	3	US-08-357-746-3
12	1298.5	32.2	906	3	US-08-630-916A-48
13	1134.5	28.1	683	3	US-08-630-916A-46
14	534.5	13.2	874	2	US-08-247-304B-8
15	534.5	13.2	874	3	US-08-767-942A-21
16	513	12.7	866	1	US-08-100-692-1
17	513	12.7	866	2	US-08-674-030-1
18	450.5	11.2	1083	3	US-08-895-601-5
19	219	5.4	486	3	US-08-348-518C-2
20	207	5.1	472	3	US-08-348-518C-5
21	207	5.1	472	3	US-08-476-509B-5
22	162.5	4.0	54	3	US-08-630-916A-124
23	158	3.9	448	3	US-08-476-509B-2
24	156.5	3.9	55	3	US-08-630-916A-75
25	154.5	3.8	454	3	US-08-348-518C-4
26	154.5	3.8	454	3	US-08-476-509B-4
27	153	3.8	224	3	US-08-630-916A-50

28 150 3.7 51 3 US-08-630-916A-117 Sequence 117, Appl  
29 147.5 3.7 54 3 US-08-630-916A-74 Sequence 74, Appl  
30 146 3.6 51 3 US-08-630-916A-73 Sequence 73, Appl  
31 145.5 3.6 58 3 US-08-630-916A-84 Sequence 84, Appl  
32 141 3.5 54 3 US-08-630-916A-118 Sequence 118, Appl  
33 140 3.5 38 3 US-08-630-916A-26 Sequence 26, Appl  
34 140 3.5 38 3 US-08-348-518C-18 Sequence 18, Appl  
35 140 3.5 38 3 US-08-476-509B-18 Sequence 18, Appl  
36 136 3.4 38 3 US-08-630-916A-36 Sequence 36, Appl  
37 135 3.3 51 3 US-08-630-916A-116 Sequence 116, Appl  
38 132 3.3 38 3 US-08-630-916A-32 Sequence 32, Appl  
39 129.5 3.2 1105 2 US-08-710-249-2 Sequence 2, Appl  
40 129.5 3.2 1105 4 US-09-220-157A-2 Sequence 2, Appl  
41 127.5 3.2 335 2 US-08-844-312-2 Sequence 2, Appl  
42 127 3.1 51 3 US-08-630-916A-115 Sequence 115, Appl  
43 124 3.1 38 3 US-08-630-916A-24 Sequence 24, Appl  
44 124 3.1 38 3 US-08-348-518C-15 Sequence 15, Appl  
45 124 3.1 38 3 US-08-476-509B-15 Sequence 15, Appl

#### ALIGNMENTS

RESULT 1  
US-08-539-205A-2  
; Sequence 2, Application US/08539205A  
; Patent No. 6001619  
; GENERAL INFORMATION:  
; APPLICANT: Beach, David H.  
; APPLICANT: Caligiuri, Maureen  
; APPLICANT: Nefsky, Bradley  
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/539,205A  
; FILING DATE: 04-OCT-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: CSV-005.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 832-1000  
; TELEFAX: (617) 832-7000  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 735 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-539-205A-2

Query Match 97.6%; Score 3939.5; DB 3; Length 735;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;

QY 1 MSNPGRRNGPVKRLTVLC AKNLVKKDFRFPDPFAKVVVDGSGGCHSTDTVXNTLDPK 60

Db 1 MSNPGRRNGPVKRLT-----GLFDPFAKVVVDGSGGCHSTDTVXNTLDPK 47

QY 61 WNCYDLYIGKSSVITISVWVHKIKKQACGLGCVRLLSNAINELKDTGTQRLDCKL 120

Db 48 WNHQYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDLCCKL 107  
QY 121 GPNNDNTVRGQIVLSQSDRIITGQGVVDCSRLFDNDLPDGWEERTASGRIOYLNHIT 180  
Db 108 GPNNDNTVRGQIVLSQSDRIITGQGVVDCSRLFDNDLPDGWEERTASGRIOYLNHIT 167  
QY 181 RTQWERTRPASEYSSPGRPISCFVDENTPISTNGATCGSSDPRLAERRVRSQRHN 240  
Db 168 RTQWERTRPASEYSSPGRPISCFVDENTPISTNGATCGSSDPRLAERRVRSQRHN 227  
QY 241 YMSRTHLHTPPDLPEGVEORTTQCGQVYFLHTQGTGVSTWHDPRVPRDLNSINCEELGFLP 300  
Db 228 YMSRTHLHTPPDLPEGVEORTTQCGQVYFLHTQGTGVSTWHDPRVPRDLNSINCEELGFLP 287  
QY 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNOLKQDQOQVVSCLPDDT 360  
Db 288 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNOLKQDQOQVVSCLPDDT 347  
QY 361 ECLTVPRYKRDVQKILRQELSOQOQAGHCRIEVSREIEFESYRQVMKRPKDLWK 420  
Db 348 ECLTVPRYKRDVQKILRQELSOQOQAGHCRIEVSREIEFESYRQVMKRPKDLWK 407  
QY 421 RLMKFRGEGLDYGGVAREWLYLSHMLNPNYGLFOYSRDDIYTLQINPDSAVNPEHL 480  
Db 408 RLMKFRGEGLDYGGVAREWLYLSHMLNPNYGLFOYSRDDIYTLQINPDSAVNPEHL 467  
QY 481 SYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND 540  
Db 468 SYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND 527  
QY 541 ITGVLDHTFCVEHNAYGEIIQHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGIEAQFLA 600  
Db 528 ITGVLDHTFCVEHNAYGEIIQHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGIEAQFLA 587  
QY 601 LQKGFNEVTPQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWPWKAV 660  
Db 588 LQKGFNEVTPQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWPWKAV 647  
QY 661 EPPDEERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 720  
Db 648 EPPDEERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 707  
QY 721 DIPPYESYKLYEKLTAIBETCGFAVE 748  
Db 708 DIPPYESYKLYEKLTAIBETCGFAVE 735

## RESULT 2

US-09-392-163A-2  
Sequence 2, Application US/09392163A  
Patent No. 6503742  
GENERAL INFORMATION:  
APPLICANT: Beach, David H.  
APPLICANT: Caliguri, Maureen  
APPLICANT: Netsky, Bradley  
TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09392,163A  
FILING DATE:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/539,205  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: CSV-005.01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 735 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-392-163A-2

Query Match 97.6%; Score 3939.5; DB 4; Length 735;  
Best Local Similarity 98.0%; Pred. No. 0;  
Matches 733; Conservative 0; Mismatches 2; Indels 13; Gaps 1;  
QY 1 MSNPGRRNGPVKRLITVLCAKNLVKDDFFRLPDPRAKVVVDGGGCHSDTDTVKNLDPK 60  
Db 1 MSNPGRRNGPVKRLITVLCAKNLVKDDFFRLPDPRAKVVVDGGGCHSDTDTVKNLDPK 47  
QY 61 WNHQYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDLCCKL 120  
Db 48 WNHQYDLYIGKSDSVTISVWNHKKIHKQAGFLGCVRLLSNAINRLKDTGYQRLDLCCKL 107  
QY 121 GPNNDNTVRGQIVLSQSDRIITGQGVVDCSRLFDNDLPDGWEERTASGRIOYLNHIT 180  
Db 108 GPNNDNTVRGQIVLSQSDRIITGQGVVDCSRLFDNDLPDGWEERTASGRIOYLNHIT 167  
QY 181 RTQWERTRPASEYSSPGRPISCFVDENTPISTNGATCGSSDPRLAERRVRSQRHN 240  
Db 168 RTQWERTRPASEYSSPGRPISCFVDENTPISTNGATCGSSDPRLAERRVRSQRHN 227  
QY 241 YMSRTHLHTPPDLPEGVEORTTQCGQVYFLHTQGTGVSTWHDPRVPRDLNSINCEELGFLP 300  
Db 228 YMSRTHLHTPPDLPEGVEORTTQCGQVYFLHTQGTGVSTWHDPRVPRDLNSINCEELGFLP 287  
QY 301 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNOLKQDQOQVVSCLPDDT 360  
Db 288 PGWEIRNTATGRVYFVDHNNRTTQFTDPRLSANLHLVLRNQNOLKQDQOQVVSCLPDDT 347  
QY 361 ECLTVPRYKRDVQKILRQELSOQOQAGHCRIEVSREIEFESYRQVMKRPKDLWK 420  
Db 348 ECLTVPRYKRDVQKILRQELSOQOQAGHCRIEVSREIEFESYRQVMKRPKDLWK 407  
QY 421 RLMKFRGEGLDYGGVAREWLYLSHMLNPNYGLFOYSRDDIYTLQINPDSAVNPEHL 480  
Db 408 RLMKFRGEGLDYGGVAREWLYLSHMLNPNYGLFOYSRDDIYTLQINPDSAVNPEHL 467  
QY 481 SYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND 540  
Db 468 SYFHFVGRIMGMAVPHGHYIDGGFTLPFYKQLLGSITLDDMELVDPDLHNSLVWILEND 527  
QY 541 ITGVLDHTFCVEHNAYGEIIQHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGIEAQFLA 600  
Db 528 ITGVLDHTFCVEHNAYGEIIQHELKPNKGSIPVNEENKKEYVRLYVNNRFLRGIEAQFLA 587  
QY 601 LQKGFNEVTPQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWPWKAV 660  
Db 588 LQKGFNEVTPQHLKTFDSEKELELIICGLGKIDVNDKVNTRLKHCTPDSNIVKWPWKAV 647  
QY 661 EPPDEERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 720  
Db 648 EPPDEERRARLLQFVTGSSRVPLOQFKALQGAAGPRLFTIHQIDACTNNLPKHAHTCFNRI 707  
QY 721 DIPPYESYKLYEKLTAIBETCGFAVE 748  
Db 708 DIPPYESYKLYEKLTAIBETCGFAVE 735

```
RESULT 3
US-08-539-205A-4
; Sequence 4, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-539-205A-4
Query Match 38.7%; Score 1564; DB 3; Length 766;
Best Local Similarity 41.1%; Pred. No. 4, 1e-138;
Matches 338; Conservative 121; Mismatches 232; Indels 132; Gaps 18;

QY 1 MNPGRNRNGPKRLTVLCAKLVKKDFRLLPDPFAKVVVDGSGQCHSTDTVRNTLDPK 60
DB 1 MNSAQSR-----RIRTVIADGLYKRDYFRPDPFAVLTVDE-QTHTTAIXKTLNPY 55
QY 61 WNOHYDLYIGKSDSVTISWNHKKTHKKGAGFLGCVRLLSNAINRLKDTGVRL--DLC 118
DB 56 WNETEVAVTNDSTAIQVFDQKF-KKKGQFLGVINLRVDVLDLAIGDEMILTRDK 114
QY 119 KUGPNDNTRVQIVVSLQ----- 137
DB 115 K--SNENTVHGKIILNSTTAQTLQVPSSAASCARTQTSITNDPQSSKSSSVSRNPA 172
QY 138 -----SDRIGTGVQVDCSLFDN-----DLPDCWEERTASGRIOVLNHTRTQ 184
DB 173 SRAGSPTRDNAPASAPASSEPRITSSPEDQVGRLLPGWERTDNLGRTYVVDHNTSTT 232
QY 185 WERPT-----RPASEYSSPCRPLSCFVDENTPIGTNGATCGQSSDPLAERRVRSQRH 238
DB 233 WIRPNLSSVAGAAAEHLSSAS--SANVTEGVQPSNNAA-----RRTEASVLT--- 279
QY 239 RNYMRTHLHPDPLPEGYEQRTOCGQVYFLHTQGVSTWHDPRVPRDLNIN----- 292
DB 280 ----SNATTAGSGELPPGWEQRYTEGRPFYVDHNTTRTTTWDPRRQYIRSYGPNPAT 335
QY 293 -----CEELGPLPCCWEIRNTATGRVYFVDHNNRTQFTDPRLSANHLVLRNQLKQD 347
DB 336 IQQPVSQLPSCGWENRLNTARVYVDHNTKTTTWDDPRLPSSL-----DQ 384

348 QQQVVSCLPDDTECLTVPRYKRDVLQKILROELSQ--QPQAGHCRIEVSREEIFEE 405
385 -----NVPQYKRDPRKLIYF---LSQPALHPLPGQCHIKVRNHFIFED 425
406 SYQVMKMRPKDLWKRLMIKFRGEEGLDYGVARWMLYLLSHEMLNPPYGLFOYSRDDIY 465
426 SYAIMEQSATDLKKRLMIKFDGEDGLDYGLSREYFFLLSHEMFNFFYCLFEYSSVDNY 485
466 TLQINPSAVNPHELSTYFHFVGRIMGMAVPHGHYIDGFTLPPFYKQLLGSITLDDMELY 525
486 TLQINPHSGINPEHLNYFKFGRVIGLAIFHRFVDAFFVVSFYKMLQKKVTLQDMESM 545
526 DPDLHNSLWILENDITGVLDHTFCVBEHNAYGEIIOHELKPNKSIIPVNEENKKEYVRLY 585
546 DAEYRSLSWLNDITGVLDLTFSEVEDNCFGEVWVTLKPNGENIEVTENKKEYVDLV 605
586 VNRFLRGIEAQLAQKGFNEVIPHLLKTTFEKELELIICGLGKIDVNDKWNKNTLKH 645
606 TVW-IQRIEEQFNAFHEGFSELTIPQELINVFDERELELLIGITSEIDMEDWKKHKDYS 664
646 CTSDSNIVKFWKAVEFEDEERRARLLQFVTGSRVPLQGFKALQGAAGPRLETHIQIDA 705
665 YSENDQIIKWFELMDWSNEKSRLLQFTTISRIPVNGFKDLQSGSDGPRKFTIERAGE 724
706 CTNNLPKHAHTCFNRIDIPPYESYEKLYEKLTLTAIETCGFAVE 748
725 -PNKLPKHAHTCFNRDLDPYTSKKDLHDHLSIAVEETIGFGQE 766

RESULT 4
US-09-392-163A-4
; Sequence 4, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-392-163A-4
Query Match 38.7%; Score 1564; DB 4; Length 766;
```

	Best Local Similarity	41.1%;	Pred. No. 4.1e-138;	
	Matches	Conservative	121;	Mismatches 232; Indels 132; Gaps. 18;
Qy	1	MSNRRRENGFVKRLRLVLCAKNVLKKDFPRLPPFAKVVVDGSGCHSTDVKNLTDPK	60	
Db	1	MSNSAQR-----RIRVIIVAADGLYKRDVRFDFDPFAVLTVDGE-QTHTTTAAIKKLTNPY	55	
Qy	61	NNQHYDLYIGKSDSVTISVMNKKHKQAGAGLGCVRLLSNAINELKDTGQRL--DLC	118	
Db	56	WNTEPEVAVNTSTAIQVFDQKXF-KKQGQGLVINLRVGDLVLAIGSGDEMLTRDUK	114	
Qy	119	KLGPNDNDTVRGQIVWSLQ-----	137	
Db	115	K--SNEVTIVGKIIINLSTTAQSLQVPSAASGARTQTSITNDPQSSKSSSVSRNPA	172	
Qy	138	-----SRDRIGTGGQVDCSRUDN-----DLDGWEERTASGRIOYLNIHTRTQ	184	
Db	173	SSRAGSPTRDNAPASPASSPRTFSFEDQYGRLLPGWERRTDNIGRTYYVDHNRTRTT	232	
Qy	185	WERPT-----RPASEYSSGCRPLSCFVDENTPISTGNGATCGSSDPLAERRVRSQRH	238	
Db	233	WTRPNLSSVAGAAAEHLSSAS--SANVTGCVQSPSSNAA-----RTEASVLT--	279	
Qy	239	RNYGRTHLHTPPDLPEGYEGRTQOQGVFLHTQTGVTWHDPRVPRDLSIN-----	292	
Db	280	-----SNATTAGSGELPPGWEQRYTPGCRPYFVDHNRTRTTTWDPDRQQYIRSYGGPNAT	335	
Qy	293	-----CEELGPLPGCWEIRNTATGRVYFVDHNRRTQFTDPRLSANLHLVLNRQNLQKQ	347	
Db	336	IQQPVSYGLPLPSGWEMLTWIRVYFVDHNTKTTWDDPRLPSSL-----DQ	384	
Qy	348	QQQVVSVCDDTECLTVPRYKRLVOKILKRLQELSQ--QPQAGHCRIEVSREIPEE	405	
Db	385	-----NVPPQKRDPRKRLIYF--LSQPALHLPQQCHI-KVRNHIPE	425	
Qy	406	SYROYMKRPKDLWKRLMKRGEGLDYGVAWEMLYLLSHMLNPYGLFOYSRDDIY	465	
Db	426	SYAEINRQATDLKRLMKDGEDGLDYGLSREYFFLLSHWFNPFCYCLEYSSVDNY	485	
Qy	466	TLQINPDSAVNPEHLSYFHFVGRINGMAVFHGYIDGGFTLPYKOLLGKSLTDDMELV	525	
Db	486	TLQINPHSGINPEHLNYFKFIRVIGLAIHRRREVDFAFFVVSFYKMLOKKVTLODMES	545	
Qy	526	DPDLHNSLVILLENDITGVLDHTICVEHNAVGEIIQHELKXNKSPVNEENKKEVRLY	585	
Db	546	DAEYRSLVILNDNITGVLDITSVDNCPGEVVTIDLAPNGRNIETVEENKREYDVLV	605	
Qy	586	VNWRFLRGIEAQLAQGFNEVIPQHLLKTTDFEKELELEIICGLGKIDVNDVNKVRLLKH	645	
Db	606	TYW-IQKRIEEQFNAFHEGFSBELIPQELINVFDSRELELLIGTISEIDMEDWKHKDYRS	664	
Qy	646	CTPDSNIVWFKVAFFDEBERREALLFOVTGSSRVPLOGFKALQGAAGRLPTTHOIDA	705	
Db	665	YSENDQIIKWFELMDENSKSKSLLOFTTGTSTRIPWNGFKLOQSGDGRKFTIEKAGE	724	
Qy	706	CTNNLPKATCTCNRIDIPPEYSYKLYEKLTAJETCGFAVE	748	
Db	725	-PNKLPKATCNRLDLPYTSKDLHKLGIAYEETIGFQOE	766	

## RESULT 5

[illegible]

QY 380 ROELSCQOQAGHCRIEVSREIPIESYQVNMKPKDLWK-RLMIKRGEGLDYGGVA 438  
 Db 557 RRLKXKQNDIPNKFEMKURARVLSDSYRIRMGVRRADFLKARLWIEFDGKGLDYGGVA 616  
 QY 439 REWLVLHSEMNPYYGLFYQSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMGMAVPHG 497  
 Db 617 REWFLISKEMNPYYGLFYQSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMGMAVPHG 676  
 QY 498 HYLDGGFTLPFYKOLLGKSIDLDDMELVDPDLHNSLVWILENDITGVLDHFTFCVEHNAYG 557  
 Db 677 KLLDGGFTLPFYKOLLGKSIDLDDMELVDPDLHNSLVWILENDITGVLDHFTFCVEHNAYG 735  
 QY 558 EIIQHELKNGKSIPIVNEENKKEYVRLVYNWFLRGIEAQLALQKGFNEVPIQHLKTF 617  
 Db 736 QTHQHELKNGKSEIIVVTKNKKKEYIYLVIQWRFVNRIOKQMAAFKEGFFELIPQDLIKIP 795  
 QY 618 DEKELELLICGLKIDVNDWKVNTRLKH-CTPDSNIVKWFKAVFFDEERARLLQVVT 676  
 Db 796 DENELELLICGLKIDVNDWKVNTRLKH-CTPDSNIVKWFKAVFFDEERARLLQVVT 855  
 QY 677 GSSRVPLOQKALOGAAGRLFTIHOIDACTNNLKAHTCFNRIDIPPYESYEKLYEXLL 736  
 Db 856 GTSRVPNGFAELYGNGPQSTVEQW-GTPEKLPRAHTCFNRIDIPPYESYEKLYEXLL 914  
 QY 737 TAIBETCGF 745  
 Db 915 MAIENTQGF 923

## RESULT 6

US-08-539-205A-6  
 ; Sequence 6, Application US/08539205A  
 ; Patent No. 6001619  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beach, David H.  
 ; APPLICANT: Caligiuri, Maureen  
 ; APPLICANT: Nefsky, Bradley  
 ; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109-2170

COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/539,205A  
 ; FILING DATE: 04-OCT-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Vincent, Matthew P.  
 ; REGISTRATION NUMBER: 36,709  
 ; REFERENCE/DOCKET NUMBER: CSV-005.01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 832-1000  
 ; TELEFAX: (617) 832-7000  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 834 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-539-205A-6

Query Match 33.5%; Score 1353; DB 3; Length 834;  
 Best Local Similarity 44.6%; Pred. No. 3.5e-118;  
 Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;

QY 159 LPOGWERRTASGRIQVNLNHTTRTOWERTRPASEYSSPG-----RPLS 203  
 Db 246 LPSGWEERKAKGRYYVYNNRNTTTTTRFIMOAEAGSGSATNSNNHIEPQIREPRS 305  
 QY 204 C---FVDENTPISGTNGATCGSQSDPRLAERRVRSQRHRYMGRTHLHTPPDLPEGYEOR 260  
 Db 306 LSSFTVLSAPLEGAQSDPVRRRAVKOTLSNPQSPQSPYNSPKPQHKVQTOSFPPGWEHR 365  
 QY 261 TTQGGQVYFHTQTGVSTWHDPRVP-----RDLSNINCEBELGLPQGWERTATGRVYF 315  
 Db 366 IAPNGRFFFTDHTKTITTTWEDPRLKFPFVHMRSKTSLNPNDLGLPQGWEEIRIHLDTFTY 425  
 QY 316 VDHNRRTTQTDPRLSANLHLVLRNQNLKDDQQQVSVSLCPDDTECLTVPR--YKSDLV 373  
 Db 426 IDHNSKITQWEDRL-----QN-----PAITGPANVPYSREPK 457  
 QY 374 QXKILQELSCQQOQPAQGHCRIEVSREIPIESYQVNMKPKDLWK-RLMIKRGEGLDYGGVA 432  
 Db 458 QKYDYFRKLLKPADIPNRFEMKLRHNNIPIESYRIRMSVKRPDVLKARLWIEFSEKGL 517  
 QY 433 DYGGVAREWLYLASHENLNPYYGLFOYSRDDIYTLQINPDSAV-NPEHLSYFHFVGRIMG 491  
 Db 518 DYGGVAREWFLSKENFNFNYYGLFEYSATDNTLTQINPNSGLCNEHDHUSYFTFGRVAG 577  
 QY 492 MAVFHGYIDGGFTLPFYKOLLGKSIDLDDMELVDPDLHNSLVWILENDITGVLDHFTFCV 551  
 Db 578 LAVFHGKLLDGGFTLPFYKOLLGKSIDLDDMELVDPDLHNSLVWILENDITGVLDHFTFCV 636  
 QY 552 EHNAYGEIIOHELKPNKGSIPVNEENKKEYVRLVYNWFLRGIEAQLALQKGFNEVPIQ 611  
 Db 637 DEENFGQTYQVLDKPNGSEIMVTNENKREYIDILVIOWRVFNRYCKQNMALLEGFTELLPI 696  
 QY 612 HLLKTFDEKELELLICGLKIDVNDWKVNTRLKH-CTPDSNIVKWFKAVFFDEERAR 669  
 Db 697 DLKIFDENELELLICGLKIDVNDWKVNTRLKH-CTPDSNIVKWFKAVFFDEERAR 755  
 QY 670 RLLOFTVTSRVPLOQKALOGAAGRLFTIHOIDACTNNLKAHTCFNRIDIPPYESYE 729  
 Db 756 RLLOFTVTSRVPNGFAELYGNGPQSTVEQW-GTPEKLPRAHTCFNRIDIPPYETFE 814  
 QY 730 KLYEKLTAETEETCGF 745  
 Db 815 DLREKLLMAVENAQGF 830

## RESULT 7

US-09-392-163A-6  
 ; Sequence 6, Application US/09392163A  
 ; Patent No. 6503742  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beach, David H.  
 ; APPLICANT: Caligiuri, Maureen  
 ; APPLICANT: Nefsky, Bradley  
 ; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
 ; STREET: One Post Office Square  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109-2170  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/392,163A  
 ; FILING DATE:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/539,205

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 834 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-392-163A-6

Query Match 33.5%; Score 1353; DB 4; Length 834;
Best Local Similarity 44.6%; Pred. No. 3.5e-118;
Matches 275; Conservative 93; Mismatches 188; Indels 60; Gaps 12;

Qy 159 LPDGEERTASGRIOYLNHITRTQWERTRPASEYSSG-----RPLS 203
Db 246 LPSGWEERKAKGRYYVYVNNHRTTTRIMQLAEDGASGATNSNNHLEIEQIRPRS 305
Qy 204 C---FVDENTPISTNGATCGSSDPLAERRVRSQHRMYSRTHLTPDPEGYEOR 260
Db 306 LSSPTVLSAPLEGAKDSPVRAVKOTLSNPFQSPQSPYNSPKQHKVQSFPLPGWEWR 365
Qy 261 TTQGGQVYFHTQGVSTWHDPRVP-----RD-SNINCEBELGLPPGWEIRNTATGRVYF 315
Db 366 IAPNGRPFIDHNTKTTTWDPRLLKPPVHMRSKTSLNPNLGLPLPGWEIRHLDKRTFY 425
Qy 316 VDHNRRTQFTDPLSANHLHLVLRNQKLDQOQQQVSLCPDDTECLTVPR--YKRDLV 373
Db 426 IDHNSKITQWEDPRL-----QR-----PAITGFAPVYSREPK 457
Qy 374 QKLKILROELSOQOQOAGHCKRIEVSREIEEESYRQVMKVRPKDLWK-RLMKFRGEEGL 432
Db 458 QKYDYPRKLUKPADIPNRFEMKLHNNIIEESYRRIMS VKRPDVLKARLWIFSESEKGL 517
Qy 433 DYGVAREWLYLHSEMLNPYGLFQYSRDDIYTLQINPDSAV-NPEHLSYFHFVRIMG 491
Db 518 DYGVAREWFLSKEMFNPFYGLFYSATDNTVTLQINPNSGLCNEHDHLSYFTIGRVAG 577
Qy 492 MAVFHGHYDGGFTLPYKOLLGKSIITLDDMELVDPDLHNSLVWILENDITGLDHTFCV 551
Db 578 LAVFHGKLDDGFTTRPFYKMLGKQIITLNDVESVDSEYNSLWILENDPT-ELDLMFCI 636
Qy 552 EHNAYGEIIQHELKPKNGKSI PVNEENKEYVYLVYVNRFLRGIEAQFLALQKGFNEVPIQ 611
Db 637 DEENFGQTYQVDLPKNGSEIMVTNENKREYIDLVIQWRFNRYQKQMNAPLEGFTELLPI 696
Qy 612 HLLKTFDEKELELLICGLKIDNDWKVNTRLK--HCTPDSNIVKFWKAVESFFDEERRA 669
Db 697 DLKIFDENELELLMCGLDGVDVNDWRQHSYKNGYC-PNHPVLOFWKRAVLMDAEKRI 755
Qy 670 RLLOFVTGSSRVPVLPQGFALQGAAGPLFTIHQIDACTNNLPAKHTCFNRIDIPPYESYE 729
Db 756 RLLOFVTGTSRVPVNGFAELVGSNGPQFLTEIOWGS-PEKLPRHTCFNRLDLPYETPE 814
Qy 730 KLYEKLTAIBETCGF 745
Db 815 DLREKLMAVENAQGF 830

RESULT 8
US-09-070-060-4
; Sequence 4, Application US/09070060
; Patent No. 5976849
; GENERAL INFORMATION:
; APPLICANT: Husted, Carolyn M.
; APPLICANT: Ghildyal, Namit
; TITLE OF INVENTION: Human E3 Ubiquitin Protein

```

```

; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHM.70312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302.886.4889
; TELEFAX: 302.886.8221
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 854 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-09-070-060-4

Query Match 33.1%; Score 1336; DB 2; Length 854;
Best Local Similarity 36.6%; Pred. No. 1.4e-116;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;

Qy 13 KRLTLVLCAL-NLVKKDPRLPDPAKVVVGSGCQCHSTVKTLPKKNQHYDLIGK 71
Db 9 QLIQTVISAKLKENKKNWFG-PSPYVEVTVD--QSKKTEKCNNTNSPKMKQPLTVTP 65
Qy 72 SDSVTISVYNNHKTIKK--QGAGFLGCVRLLSNAINRLKDTGQRLDLCKLGPNDNDTVR 129
Db 66 TSKLCFRVWSHQTLKSDVLLGTAGLDIYETLKSNNMKLEEV---VMTLQLVGDKEPTETM 122
Qy 130 GQIVY---SLQSRDRIGTGGQV-----DCSRLFDNLPD-----GWEE 165
Db 123 GDLSCVLDGLQVEAEVNTGETSCSESTQDNDGCRTRDDTRVSTNGSEDEVAASGENK 182
Qy 166 RTASGRYCYLNHITRTTQWERPTRPASEYSSGPRPLSCFVDENT-----PIS 213
Db 183 RANGNNSPSLNGGPKSRPPRPPRPPP--PTPRPASVNGSPSTNSDSSTGSLPPT 240
Qy 214 GTN-----GATCG-----QSSDPLAERRVRSQHRNYM-----242
Db 241 NINVTNTSEGATSLIIPLTISGSGPRPLNTVSQAPLPGEQRV-DQGRVYVVDHV 299
Qy 243 -SRTHLHTPPDLPEGYEORTTQGOVYELHTQTGVSTWHDPRVP-----285
Db 300 EKRTTDRPEPLPGWEERRVDNMGRYYVDHFTTTTQWRPTLSVRNYEQWLQRSOLQ 359
Qy 286 -----RDL-----SNINCEBELGLPPGWEIRNTATGRVYFVDHNNRTQFTD 327
Db 360 GAWQFNQRFYVNGQDLFATSONKEFDPLGPLPGWEKRTDSNGRVYFVNHTITOWED 419
Qy 328 PRLSANHLVLRNQKLDQOQQQVVSLLCP---DDTECLTV---PR-----Y 368
Db 420 PRSQGQ----LNEKPLPEGWEWRFTVDGIPYFVDHNRRTATYIDPRTOKSALDNGPQIAY 475

```

```

QY 369 KRDVOKILR---OELSOQOQACHCRIEVSREIPEESYQVMMKPKDLWKRLMIK 425
Db 476 VRDFKAKVQVFRWCQOLAMPQ---HIKITVTRKTLFEDSPQOIMFSFQDLRLRLWVI 531
QY 426 FRGEGLDYGGVAREWLYLLSHEMLNPPYXGLFOYSRDD:YTLQINPDSAVNPEHLSYFHP 485
Db 532 FPGEGLDYGGVAREWFFLLSHEVLNPMYCLFAYAGKDNVCLQINPASYINPDHLKYFF 591
QY 486 VGRIMGMAVPHGYIDGGFTLPYKOLLGKSIITLDMELVDPDLNLSLWILENDITGV- 544
Db 592 IGRFIAMALFHGFIITGSLPPIKILNKPGLKOLLESIDPEFYNSLIWVKNNEIECG 651
QY 545 LDHTFCVHNAYGEIICHKLKPKNGKSIPIVNEENKKEYVRLYNWRLRGLRGLAEQFLALOKG 604
Db 652 LEMYFSVDKEILGEIKSHDLKPNNGNLTVEENKEEYIRMAEWRLSRGVEEQTAQAFEG 711
QY 605 FNEVIPHLLKTFDEKELELIICGLGKIDVNDKWNTRKHKCTPDSNIVKMFKAVEFFD 664
Db 712 FNEILPQOYLQVFDPAKELEVLICGMQEIIDLNDQWQHAIYRHYTRTSKQIMWFQVKEID 771
QY 665 EERRARLLOFVTGSSRVLPQGFALOGAAGPRLFTTHQIDACTNNLPKAHTCFNRIDIPP 724
Db 772 NEKRMELLOFVTGTCRLPVGGFADLMGSGNGPKFCIEKVGK-ENWLPKSHTCFNRDLDP 830
QY 725 YESYKLYEKLTLTAIBETCGFAVE 748
Db 831 YKSYEQLEKLLFAIBETEGFQGE 854

RESULT 9
US-09-357-746-4
; Sequence 4, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM 70312 N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-357-746-4

Query Match
Best Local Similarity 33.1%; Score 1336; DB 3; Length 854;
Matches 316; Conservative 130; Mismatches 272; Indels 146; Gaps 24;

QY 13 KLRLTVLCAK-NLVKDKDFRLPDPFAKVVVGGSGOCHSTDIVKNTLDKKNQHDLYIGK 71
Db 9 QLOITVISAUKENKXNWF-GPSPVETVTD--GQSKTEKCNNTNSPKWQPIITVITP 65
QY 72 SDSVTISVNNKHKKHK--QGAGFLGCVRLLSNAINRLKDTGYQRLDCLKLPNDNITVR 129
Db 66 TSKLCFRVWASHOTLKSDVLLGTAGLDIVETLKSNNMKLEEV--VMTLQLVGDKPETM 122
QY 130 GOIVV---SLQSRDRIGTGGQV-----DCSLRFDNLDLP-----GWEE 165
Db 123 GDLVCLDGLQVAEVVTNGETSCSESTQNDGCGTRDTRVSTNGSEDPEVAASGENK 182
QY 166 RRTASGR:QYLNHITRITQWERTPASEYSPGRPSCFVDENT-----PIS 213
Db 183 RANGNNSPSLGGFKPSRPPRPSP--FTPRPASVSPSTNSDGSSTGSLDPT 240
QY 214 GTN-----GATCG-----QSSDPLAERRVRSQHRNYM----- 242

```

```

Db 241 NTNNTSTSEGAISGLIIPLTISOGSGRPLNTVSOAPLPFGWEQRV--DOHGRVYVDHV 299
QY 243 -SRTHLHTPPDLPGYEQRITTOQGVYFLHTQTGVTGVTWHDRVP----- 285
Db 300 EKRTTWDREPLPFGWEERVDNMGRIYVDHFTTTTQWRPTLESVRNVEQWQLORSOLQ 359
QY 286 -----RDL-----SNINCEELGPLPPGWEIRNTATGRVYFVDNNHTTQPTD 327
Db 360 GAOQFNORFIYGNQDLFATSONKEFDFPLGLPFGWEKRTDSDNGRVYFVNHTRITQWED 419
QY 328 PRLSANLHLNQRNQLKQDQOQVWSLCP---DTECLTV---PR-----Y 368
Db 420 PRSQOQ---LNEKPLPEGWEMRFTVGIPIFYVDHNRATYIDPRICKSALDNGPQIAY 475
QY 369 KRDVOKILR---OELSOQOQACHCRIEVSREIPEESYQVMMKPKDLWKRLMIK 425
Db 476 VRDFKAKVQVFRWCQOLAMPQ---HIKITVTRKTLFEDSPQOIMFSFQDLRLRLWVI 531
QY 426 FRGEGLDYGGVAREWLYLLSHEMLNPPYXGLFOYSRDDIYTLQINPDSAVNPEHLSYFHP 485
Db 532 FPGEGLDYGGVAREWFFLLSHEVLNPMYCLFAYAGKDNVCLQINPASYINPDHLKYFF 591
QY 486 VGRIMGMAVPHGYIDGGFTLPYKOLLGKSIITLDMELVDPDLNLSLWILENDITGV- 544
Db 592 IGRFIAMALFHGFIITGSLPPIKILNKPGLKOLLESIDPEFYNSLIWVKNNEIECG 651
QY 545 LDHTFCVHNAYGEIICHKLKPKNGKSIPIVNEENKKEYVRLYNWRLRGLRGLAEQFLALOKG 604
Db 652 LEMYFSVDKEILGEIKSHDLKPNNGNLTVEENKEEYIRMAEWRLSRGVEEQTAQAFEG 711
QY 605 FNEVIPHLLKTFDEKELELIICGLGKIDVNDKWNTRKHKCTPDSNIVKMFKAVEFFD 664
Db 712 FNEILPQOYLQVFDPAKELEVLICGMQEIIDLNDQWQHAIYRHYTRTSKQIMWFQVKEID 771
QY 665 EERRARLLOFVTGSSRVLPQGFALOGAAGPRLFTTHQIDACTNNLPKAHTCFNRIDIPP 724
Db 772 NEKRMELLOFVTGTCRLPVGGFADLMGSGNGPKFCIEKVGK-ENWLPKSHTCFNRDLDP 830
QY 725 YESYKLYEKLTLTAIBETCGFAVE 748
Db 831 YKSYEQLEKLLFAIBETEGFQGE 854

RESULT 10
US-09-070-060-3
; Sequence 3, Application US/09070060
; Patent No. 5976849
; GENERAL INFORMATION:
; APPLICANT: Huestad, Carolyn M.
; APPLICANT: Chidyal, Namit
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998

```



Db 458 PRTGKS---ALDNGPOI-----AYVRDFKAKVQYFRFWCQQLA 492  
QY 385 OQPOAGHCRIEVSREIEFESYVOMKMRPKDLWKRLMKRGEGLDYGGVAREWVYL 444  
Db 493 MPQ-----HIKIVTRKTLFEDSFQIMSPSQDLRRLLWVIFPGEGLDYGGVAREWVFL 548  
QY 445 LSHMLNPPYGLFOYSRDDIYIQLNPDSDAVNPEHLSYFHVGRIMGMAVPHGHYIDGGP 504  
Db 549 LSHVLNPNMYCLFVAGKDNVCLQINPASYINPDHLKYFRFGFTAMALPHKGFIDTGF 608  
QY 505 TLFPYKOLGKGTITLDDMELVDPDLHNSLVWLTLENDITGV-LDHTFCVEHNAAYGIIIOHE 563  
Db 609 SUPFYKRIILNKVGLJDESIDPEFNSLIWKENNIEBCDLEMVFSVDKEILGSIKSHD 668  
QY 564 LKPNKSIPIVNEENKKEYVRLVYNNRFLRGLGIAQFLALQKGFNEVLPQELLKTFDEKELE 623  
Db 669 LKPNNGNIIIVTENKEEYIRMAEWRLSRGVEEQTAFFEGFNEILPQOYLQYFPAKELE 728  
QY 624 LIICGLGKIDVNDKVNTELKCTPDSNIVKFWKAVEFFDEERARLLQFVTGSSRYPL 683  
Db 729 VILCGMQLDNDWGRHAIYRYARTSKQIMWFQFVKEIDNEKRWLLQFVTGTCRLPV 788  
QY 684 QGFALQAGAPRLFTIHOIDACTNNLPKHAHTCFNRIDIPPYESYEKLYEKLLTAIBETC 743  
Db 789 GGFADLMGNGKQFCIEKVGK-ENWLPKSHTCFNRDLDPYKSYEQLEKLLPAIBETE 847  
QY 744 GRAVE 748  
Db 848 GFGQE 852

## RESULT 12

US-08-630-916A-48  
; Sequence 48, Application US/08630916A  
; Patent No. 601137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: United States  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/630,916A  
; FILING DATE: 03-APR-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MISROCK, S. LESLIE  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-203  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 896-8864/9741  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 906 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-630-916A-48

## Query Match

Best Local Similarity 32.2%; Score 1298.5; DB 3; Length 906;  
Matches 271; Conservative 101; Mismatches 216; Indels 57; Gaps 9;  
QY 122 PNDNDTVRGQIVSVLQSRDRIGTGGVQVDCSRULFNDNDLPDGMWEERRTASGRIOYLNHTR 181  
Db 301 PNPNTTSLPAPATPAEGEPEPSTGTOQLPAAQAAPDALPAGWEQRELPNGRVVYVDHNTK 360  
QY 182 TTOWERTPRASEYSSPGRPLSCFVDENTPISGTNGATCGSSDPRLAERRVRSORHN- 240  
Db 361 TTTWERPLPFGWEKRTDPRGRFYVDHNT-----RTTWRPTAEYVRVYEQWSORNO 414  
QY 241 -----YMSRTHLHTP-PDLPEGYEQRTOQGOVYFLHTQTGVSTWHDPRV 284  
Db 415 LQAMQHFSQRFLYQFWSASTDHLPLPPEWKEK-QDNGRVVYVNNHTRTTQWEDPRT 473  
QY 285 PRDLNINCELGLPLPGWEIRNATGRVYVDHNNRTTOFTDPRLSANLEHLVLRNQQL 344  
Db 474 QGMIQE-----PALPPGEMWKYTSBGVRYVDHNTTTTTFKDPRPGE----- 516  
QY 345 XDQOQOQVWSLCPDDTTECLTVPRYKRDVLQKILRQLSQQQPQAGHCRIEVSREIFE 404  
Db 517 -----SGTKQSGPGAYDRSFRWKYHQFR-FLCHSNALPSHVKISVSRQTLFE 562  
QY 405 ESYRQWKMPPKDLWKELMKFRGEBGLDYGGVAREWLYLLSHMLNPPYGLFOYSRDDI 464  
Db 563 DSFOQIMNMKPYDLRRRLYIIMEGEGLDYGGIAREWFFLLSHEVLNPNMYCLFEYAKNN 622  
QY 465 YTLQINPDSAVNPEHLSYFHVGRIMGMAVPHGHYIDGGFTLFPYKOLLGKSTITLDDMEL 524  
Db 623 YCLOINPASSINPDHLTYFRIGFTAMALYHKGFDITGFTLFPYKMLNKRFTLKDL 682  
QY 525 VDPDLHNSLVILENDITGV-LDHTFCVEHNAVGEIIOHELKPNKGSIPVNEENKKEYVR 583  
Db 683 IDEFEYNSIWMKENNLEECGLFLYFQDMLEILGKVTHELKEGGSIRVTEENKKEYIM 742  
QY 584 LYNNWRELGRIEAQFLALQKGFNEVLPQELLKTFDEKELELIICGLGKIDVNDKVNTRL 643  
Db 743 LLTDWRTREVEBQTKAFLDGFNEVAPLEWLRVDFDEKELELMCGMQLDNDWGRHAI 802  
QY 644 KHCTPDSNIVKFWKAVEFFDEERRARLLQFVTGSSRYPLQGFALQAGAPRLFTIHOI 703  
Db 803 RHVTKNSKQIQWFQVQVVKEMDNKRIQLQFVTGTCRLPVGGFAELIGSNPGPKFCIDKV 862  
QY 704 DACTNNLPKHAHTCFNRIDIPPYESYEKLYEKLLTAIBETCGFAVE 748  
Db 863 GKET-WLPRSHTCFNRDLDPYKSYEQLEKLLYAIETEGFQGE 906

## RESULT 13

US-08-630-916A-46  
; Sequence 46, Application US/08630916A  
; Patent No. 601137  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; APPLICANT: Kay, Brian K.  
; APPLICANT: Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME  
; NUMBER OF SEQUENCES: 124  
; CORRESPONDENCE ADDRESS:  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: United States  
; COUNTRY: United States  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 683 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-630-916A-46

Query Match      28.1%; Score 1134.5; DB 3; Length 683;
Best Local Similarity 43.7%; Pred. No. 9.4e-98;
Matches 241; Conservative 84; Mismatches 182; Indels 45; Gaps 11;

QY 159 LPDGEERRTAGSRIQYLNHITRTQWER--PTRPASEYSSPGRPISCFVDENPTISGYN 216
DB 162 LPSCWEQRKDPHGRTYVYDHNTRTTTWERPQPLPGWERRVDDRRYVYVDHNTRTTTQ 221
QY 217 GATC-----GQSDPLRAERRVRSQRHNYM--SRTHLHTPPD---LPEGEQRTTQ 265
DB 222 RPTMESVRNFEQWQSQNRQJGAMQFNQRYLYSASMLAENDPYGLPPGWEXKRVDS 281
QY 266 QVYELHTQTGVTWHDRVPRDLSINCEBELGPLPGCWEIRNTATGRVYFVDHNNRTTQ 325
DB 282 RVYFVNNTKTQWEDRT--QGLQN---EE--PLPEGWEIRYTRGVYFVDHNTTTF 335
QY 326 TDPRLSANLHLVLRNQLKDDQQQVVSCLPDDTECLTVPRYKRDVLQKILKILQELSQ 385
DB 336 KDPK-----NGKSVTKGSPQIA-----YERGFWRKLAHRY--LCQ 370
QY 386 QPQAGHCRIVSEETIFESYRVQWVRPKDLWKRLMKFRGEEGLDYGVAWEMLYLL 445
DB 371 SNALPSHVKNVSRQTLFESFQIMALKYDLRRLYVIFRGEGLDYGGLAREWFFLL 430
QY 446 SHEMLNPYGLFQYSRDDIYTLQNPDSAVNPEHLSYFHFVGRIMGNVAPHYIDGGFT 505
DB 431 SHEVLNPMYCLFEYAGKNNYCLQINPASTINPDHLSYFCIGRFIAPALPHGKFIDTGS 490
QY 506 LPFYKQLIGKSLTDDMELVDPDLHNSLVMTLENDITGV--LDHTFCVEHNAIGEIIQHEL 564
DB 491 LPFYKRLSKKLTIKLESIDTEFYNSLIWRDNIEECGLEMYFSVDMILKVTSHDL 550
QY 565 KPNKSIIPVNEENKEYRYLVYVNRFLRGIEAQLAKQGFNEVYIPOHLLKTTDEKELEL 624
DB 551 KLGSSNILLVTEENKDEYIGLMTWERSKGVQEQTKAFLDGFNEVWPLWLQYFDEKELEV 610
QY 625 IICGLGKIDVDNWKVNTRLKCHTDPDSNIVKWFKAZFFDEERRARLLOFVTCSSVPLQ 684
DB 611 MLCGMQEVLDADQWRVTYRHYTRNSKQIIFWFOFKETDNEVRMLULQFVTGTCRLPLG 670
QY 685 GFALQGAAGPR 696
DB 671 GFALMSGNGPR 682

```

## RESULT 14

```

US-08-247-904B-8
; Sequence 8, Application US/08247904B
; Patent No. 5981699
; GENERAL INFORMATION:
; APPLICANT: Rolfe, Mark
; APPLICANT: Eckstein, Jens W.
; APPLICANT: Draetta, Giulio

```

```

; TITLE OF INVENTION: Human Ubiquitin Conjugating Enzyme
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/247,904B
; FILING DATE: 23-MAY-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-029.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 874 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-247-904B-8

Query Match      13.2%; Score 534.5; DB 2; Length 874;
Best Local Similarity 25.1%; Pred. No. 4.5e-41;
Matches 185; Conservative 117; Mismatches 275; Indels 161; Gaps 23;

QY 105 NKLKDTGQRDLDCVGN---DNDTVRGQIVVLSQSRDRIGTG-----GQVVVDCS 152
DB 202 SRIGSSQGDNNLQKGLDGDVSDVIDAIR--RVYTELLSNEKIEIAFLNALVYLSPNVECD 260
QY 153 RLFDNDLPDGEERRTAGSRIQYLNHITRTQWERPTRPASEYSSPGRPISCFVDENPTI 212
DB 261 LTYHN-----VYSEDENLNLF--IIGMENRNLHSPEYLEMALPLFCRMSKLPL 308
QY 213 SGTNG-----ATCGQSSDPLRAERRVRSQRHNYMSRTHL----- 247
DB 309 AAQKGLIBLWSKNADQIRRMWETFOQLITTKVISNBFNSRNLVNEFSRNLVNDDDAIV 368
QY 248 -----HTPPD-----LPS-----GYEORTTQQQVYVFLH 271
DB 369 AASKCLKMVYANVVGGEVDTHNNEEDDEEPIPESSSELTLQELLGEERRNKKGLRVDPLE 428
QY 272 TGTGVTWHDRVPRDLSINCEBELGPLPGCWEIRNTATGRV-----YFVDHNNRT 322
DB 429 TELGVKT-----LDCKK--PLIPFEFINEPLNEVLENDKDYTFKVFETENKF 474
QY 323 TQFTDPRL--SANLHLVLRNQLKDDQQQVVSCLPDDTECLTVPRYKRDVLQKILR 380
DB 475 SMTCPFFILNAVTKNLGLYYDNRIMYSERITVL----- 509
QY 381 QELSQQQPQAGHCRIVSEETIFES-----YQVWVRPKDLWKRLMKFRGEEGLDYGCV 437
DB 510 YSLVQGGQQLNPYLRLKVRDRHIIIDALVRLEMIAMENPADLKQLYVEFEQGGVDEGV 569
QY 438 AEEMLYLLSHEMLNPYGLFQYSRDDIYTLQNPDSAVNPEHLSYFHFVGRIMGNVAPHG 497
DB 570 SKEFFQLVVEELFNPDIGNFTYD--ESTKLFWNFESSF---ETEGQFTLIGIVGLAIYN 625
QY 498 HYIDGGFTLPFYKQLIGKSLTDDMELVDPDLHNSLVMTLE--NDITGLVDHTTCV--EHN 554
DB 626 CILDVHFPMVYVYKLMGKKGLFVLDGLDGHVLYQSLKDLLEYVGNVEDDMMITFIQSQTN 685

```

QY 555 AYGEIIQHELKPKNGKSIPIVNEENKKEYVPLVYVNWRLRGIEBAQFLALQKGFNEVIPPQHLL 614  
DB 686 LFGNPMYDLKENGDKIPITNENKREKFEVNLVSDYILNKSEVKOFKAFRRGFHMVNEGSL 745  
QY 615 K-TFDEKELELIICGLGKIDVNDWKVNTLX-HCTPDSNIVKFWKAVEFFDEERRALL 672  
DB 746 KYLFRPEEIELLIGSRNLDFOALEBTTEYDGGYTRDSVLIREFWEIVHSFTDEOKRLFL 805  
QY 673 QFTGSSRVPLQ---FKALQGAAGPRLFTIHQIDACTNNLPKATCTNRIDIPYESYE 729  
DB 806 QFTTGTDRAPVGGGLGKLMIAKNGPD-----TERLPTSHTCFNVLLLPYSSKE 855  
QY 730 KLYEKLLTAIEETCGFAV 747  
DB 856 KLERLLKAITYAKGFGM 873

RESULT 15  
US-08-767-942A-21  
Sequence 21, Application US/08767942A  
Patent No. 6068982  
GENERAL INFORMATION:  
APPLICANT: Rolfe, Mark  
APPLICANT: Chiu, M. Isabel  
APPLICANT: Berlin, Vivian  
APPLICANT: Damagnez, Veronique  
APPLICANT: Draetta, Giulio  
APPLICANT: Guillaume, Cottarel  
TITLE OF INVENTION: UBIQUITIN CONJUGATING ENZYMES  
NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/767,942A  
FILING DATE: 17-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MIV-029.04  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 874 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-767-942A-21

Query Match 13.2%; Score 534.5; DB 3; Length 874;  
Best Local Similarity 25.1%; Pred. No. 4 5e-41;  
Matches 195; Conservative 117; Mismatches 275; Indels 161; Gaps 23;  
QY 105 NPLKDTGYORLDCKLGNP-----DNDTVRGQIVWSLQSRDRIGTG-----GOVYDCS 152  
DB 202 SRIGSSQGNLNLQKLPDSDVSDIDAIR-RVYTRLLSNEKIETAFNLVYLSNVNVEDC 260  
QY 153 RLFDNDLPDGEERRPASPQIOYLNHIITTTQWERPTPASEYSSGPRPLSCFVDENTPI 212  
DB 261 LTYHN-----VYSDPNLNLFP--IIGWENRNLSPEYLEMALPLFCCKAMSKLPL 308  
QY 213 SCTNG-----ATCQSSDPRLAERRVRSQFHRNMYGRTHL----- 247

DB 309 AAQKGLRLMSKYNADQIRRMETFOQLITYKVIISNFNSRLNVFNFSRLNVDNDAIV 368  
QY 248 -----HTPPD-----LPE-----GYEQRITQQOQVYFLH 271  
DB 369 AASKCLMWVYVYANVVGGEVDTNHNEEDDEBPIDESSLTI-QELLGEBERRNKKGLRVDPLE 428  
QY 272 TOTGVSTWHDPRVPRDLSNINCEBELGPLPPGWEIRNATGRV-----YFVDHNNRT 322  
DB 429 TELGVKT-----LDCRX--ELIPFEFINEPLNEVLEMDXDTFFFKVETENKF 474  
QY 323 TQFTDPRPL--SANLHLVLRQNLKDOQOQVSVLCPDDTECLTVPRYKRDVLQKILR 380  
DB 475 SFMTCPEILNAVTKNLGLYYDNIRMYSERITVL----- 509  
QY 381 QELSQQOQPOAGHCHRIEVSREEIFEES---YRQVNMKRPKDLWKPLMKPKRGEGLDYGCV 437  
DB 510 YSLVQGGQLNPFYLRKLVRRDHIIDALVRLEMAMENPADLKKQLYVEFEQEGQVDSGGV 569  
QY 438 AREWLVLLSHEMLNPPYGLFCYSRDDIYTIQINPDSAVNPEHLSYFHFVGRIMGMAVPHG 497  
DB 570 SKEFFQLVVEEIFNPDIGMFTYD-ESTKLPWFNPSSF---ETEGQFTLIGIVLGLAIYNN 625  
QY 498 HYIDGGFTLPFYKOLLGKSITLDDMELVDPDLHNSLVWILE--NDITGVLDHTFCV-EHN 554  
DB 626 CILDVHFPMVYVRKLMGKGLFVDLGDSPHVLQSLKDLLEYVGVNVEDDMWITFQISQTN 685  
QY 555 AYGEIIQHELKPKNGKSIPIVNEENKKEYVPLVYVNWRLRGIEBAQFLALQKGFNEVIPPQHLL 614  
DB 686 LFGNPMYDLKENGDKIPITNENKREKFEVNLVSDYILNKSEVKOFKAFRRGFHMVNEGSL 745  
QY 615 K-TFDEKELELIICGLGKIDVNDWKVNTLX-HCTPDSNIVKFWKAVEFFDEERRALL 672  
DB 746 KYLFRPEEIELLIGSRNLDFOALEBTTEYDGGYTRDSVLIREFWEIVHSFTDEOKRLFL 805  
QY 673 QFTGSSRVPLQ---FKALQGAAGPRLFTIHQIDACTNNLPKATCTNRIDIPYESYE 729  
DB 806 QFTTGTDRAPVGGGLGKLMIAKNGPD-----TERLPTSHTCFNVLLLPYSSKE 855  
QY 730 KLYEKLLTAIEETCGFAV 747  
DB 856 KLERLLKAITYAKGFGM 873

Search completed: September 21, 2004, 07:49:22  
Job time : 28.8994 secs

This Page Blank (uspio)